DOCUMENT RESUME

ED 123 240 95 TM 005 099

AUTHOR Novick, Melvin R.
TITLE New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction. Final Report.
INSTITUTION American Coll. Testing Program, Iowa City, Iowa.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-2-0067
PUB DATE 21 Dec 73
GRANT OEG-0-72-0711
NOTE 506p.

EDRS PRICE MF-$1.00 HC-$27.45 Plus Postage
DESCRIPTORS *Bayesian Statistics; Computer Oriented Programs; Computer Programs; *Criterion Referenced Tests; *Decision Making; *Individualized Instruction; Probability; Program Development; Programing Languages; *Statistical Analysis; Student Testing
IDENTIFIERS Individually Prescribed Instruction; *Test Length

ABSTRACT

This project is concerned with the development and implementation of some new statistical techniques that will facilitate a continuing input of information about the student to the instructional manager so that individualization of instruction can be managed effectively. The source of this informational input is typically a short criterion-referenced test specific to one or more behavioral objectives being taught. The fundamental issue is one of test length. Part I of this project is a brief and broad structuring of the problem and work that has been done to solve it. Part II consists of a broad overview of the project in which each prepared paper is concisely stated. Part III discusses these papers and the new techniques and methods provided by each paper. The papers are included in four appendices; some are not available here due to copyrights but a journal reference is given. The four papers in Appendix I are directly concerned with the implementation of new methods within individually Prescribed Instruction. Appendix II deals with possible further developments in decision-theory application. Appendix III treats the core-mathematical theory underlying the proposed applications, and Appendix IV deals with the computer problems involved in such applications. (RC)

TM

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

THE AMERICAN COLLEGE TESTING PROGRAM

December 21, 1973

Final Reports Unit
Research Analysis and Allocation Staff
Bureau of Research
U.S. Office of Education
Washington, D. C. 20202

Dear Sir:

Attached is our final report to OE Grant No. 0-72-0711 on "New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction".

Sincerely,

Melvin R. Novick
Director, Psychometric Research
Research and Development Division

James McAdorough
Signature of Contract Officer

Signature of Principal Investigator

December 27, 1973
Date

December 21, 1973
Date

National Office P.O. BOX 168, IOWA CITY, IOWA 52240 / PHONE (319) 351-4470

I. INTRODUCTION

The individualization of instruction throughout the American educational system continues to increase at an accelerating pace, reflecting an increase in concern that every student receives the maximum potential benefit to be derived from that system. The very nature of that individualization requires a continuing input of information about the student to the instructional manager so that the individualization can be managed effectively. The source of that informational input is typically a short criterion-referenced test, specific to one or more behavioral objectives being taught in the particular training module that the student is currently undertaking. The project upon which we are reporting here is one that has been concerned with the development and implementation of some new statistical techniques, that make it possible to provide this informational input in a more efficient manner. The fundamental issue is one of test length. Every minute devoted to testing is now a minute less that is devoted to instruction. Thus there is a continuing desire to minimize the time allocated to testing in each instructional unit.

This report comes in seven parts. Part I, this Introduction, is a very brief and broad structuring of the problem and work that has been done to solve this problem. Part II consists of a broad overview of this project in which the contribution of each paper prepared for the project is stated in concise form so that the reader will have available a broad relief map of the territory to be explored. This overview is then flushed out in Part III with an integrated discussion of these papers and a more detailed indication and discussion of the new techniques and methods provided by each paper. Throughout this discussion we attempt to keep the technical language at a most moderate level, even at the expense at times of ignoring some technical difficulties and detail. It is our hope that this section will be readable, not only by specialists in educational measurement and statistics, but by the broader range of people who are directly concerned with the instructional-decision process. Following this discussion, the technical developments made on this project are presented in detail in four appendixes, each focused around one major aspect of the instructional decision problem. The four headings for these appendixes are discussed and integrated as part of the discussion in this introduction. Those concerned with the actual implementation of the instructional decision procedure will need to study these technical papers in great detail.

It is generally recognized that, taken on their own, scores from short tests do not provide adequately definitive information about student competency, so that sufficiently accurate instructional decisions can be made consistently. On the other hand, there is enormous interest within the group of instructional managers to keep

testing time low and specifically to keep test lengths below the 20-item mark, and preferably within the 6 to 12-item range. If this desire is to be satisfied, and at the same time we are to assure ourselves that accurate decisions are being made, on the average, then most of the time there will be a need to bring some further information into the decision-making process.

There is no difficulty in recognizing that such information exists, in the background, and is available for use. The difficulty is one of finding the technology for quantifying this information and incorporating it into the decision-making process. One of the things that we know about individualized instructional programs, for example the University of Pittsburgh's Individually Prescribed Instructional (IPI) Program, is that students take very short curriculum embedded tests within each module, and continue to receive instruction until such time as there is good reason to believe that they can pass the end of module posttest. Thus the very nature of the IPI module is such as to reduce the variability of posttest performance levels, and to suggest that there is reasonable prior probability that any particular student will pass the posttest. It is precisely this information which can and needs to be incorporated in the decision analysis which has a short criterion-referenced posttest as its direct data input.

These considerations clearly suggest that Bayesian methods which combine prior (that is, background) information with direct observational information, may be useful in sharpening IPI decision making. We shall therefore turn to the technical problems of quantifying this background information for a Bayesian analysis. If IPI programs were uniformly administered throughout the country, it would be possible to gather background information at various locations and thus to construct data-based prior distributions for adoption at all IPI installations. Unfortunately, from the decision making point of view, but not necessarily from the educational point of view, IPI methods are administered with substantial local flexibility. As a result, data gathered from one installation may not necessarily be relevant to the implementation of IPI at another installation.

At a particular IPI installation at which the instructional process has stabilized and made uniform and where historical records are available, it will be possible to use standard Bayesian methods to develop an instructional decision making process. However, as a fast growing and evolving entity, IPI is finding itself being modified continuously in current schools and being introduced into new schools where such background information has not been gathered, and therefore standard Bayesian techniques cannot be used with any expectation of guaranteeing high decision-making accuracy. It therefore becomes necessary to devise and implement new Bayesian methods that make it possible to simultaneously gather the background information necessary to provide prior information and at the same time, gather and integrate the direct information concerning the performance on each individual student. It is this requirement for simultaneous data gathering that demanded the development of new and

complex statistical methods, that make it possible to simultaneously incorporate both direct observations and collateral group information.

A gross statistical methodology for accomplishing this has been available in embryonic form since the early days of Truman Kelley's Statistical Methods text (1923). Kelley concerned himself with the estimation of true score for a group of individuals and showed, by a standard application of regression theory, that improved estimates of true score could be obtained through a simultaneous estimation procedure, based both on the direct observations of the particular individual and estimates of the group mean and true score variation across individuals, all made simultaneously. The celebrated Kelley formula is of a form which estimates each person's true score as a weighted average of his observed score, and the mean of the observed scores throughout the population. Thus, each individual's true score estimate is regressed from his observed score towards the overall mean observed score in the population. A specific formula was given by Kelley which showed the extent to which improvement in estimation could be accomplished using this technique. In 1956, Charles Stein used a similar kind of logic to show that in most situations, when a group of individual parameters are being estimated simultaneously, the standard estimates, which here correspond to the mean scores for each of the individuals, are inadmissible in a strict statistical sense in that there are always available estimates with better mean-squared error. Stein proposed a class of estimates of a form remarkably similar to those given by Kelley. In fact they differ from Kelley's estimates only to the extent that the regression to the mean is statistically somewhat less than with the Kelley estimates. Stein's work opened up the whole field of simultaneous estimation, and his ideas stimulated similar developments from other classical, from Bayesian, and from empirical Bayesian points of view. From a Bayesian point of view the first comment on this possibility was made by Lindley (1962) in the discussion at the Royal Statistical Society of a second paper by Stein (1962). A paper by Box and Tiao (1968) on the Bayesian Estimation of Means for the Random Effects Model, provided a specific Bayesian methodology, though Box and Tiao did not directly confront the simultaneous estimation problem. Recently this approach has been taken up in great detail by Lindley and his associates, and a general formulation has been provided in a paper by Lindley and Smith (1972). The possibility of application of these methods in educational research was noted several years earlier by Novick (1970). Building upon the work of Lindley (1971), Box and Tiao (1968), Stein (1956, 1962), Kelley (1923), and others, Novick, Lewis, and Jackson (1973) specialized the Bayesian simultaneous estimation procedure to the problem of the estimation of proportions in m-units and shortly after doing this, as indicated in that paper, it became clear that this method might have useful application in the field of criterion-referenced tests. Stimulated by discussions with Ronald Hambleton, who was an ACT postdoctoral fellow during the Summer of 1971, a proposal was forwarded to the Office of Education for the funding of a project to further develop and tailor these Bayesian methods for

potential implementation within IPI. In this proposal it was pointed out that, in theory, these methods could increase the effective-test length the equivalent of 6 to 25 observations through the use of collateral information. To put it the other way around, this allows a reduction in the required test length anywhere from 6 to 25 items while maintaining the same level of precision.

It was also noted in the proposal that the primary new statistical development would focus on a shift in strategy from the point estimation of individual ability levels, or true performance levels, to the determination of the probability that a particular individual mastery level is larger than some specified criterion level. From an educational point of view, this represented a tailoring of the theory to criterion-referenced testing rather than norm-referenced testing. Statistically, this meant that the output of the Bayesian analysis would not be a joint-point (modal) estimate of ability scores for students, but rather for each student, the aposteriori determination of the probability that his mastery score is larger than some specified criterion level. The strategy as noted in the proposal would be to use these aposteriori probabilities in a standard Bayesian decision theoretic context for deciding whether or not an individual student should be advanced to the new unit of instruction or retained for further work in the current module. In the proposal it was suggested that the simplest reasonable approximation to reality would be to assume a threshold loss function which specified zero losses for correct positive and negative decisions, and losses a and b, respectively, $(a, b > 0)$, for false positive and false negative decisions. For the most part this loss structure is assumed throughout the work on this project though in one paper we do indicate that this is only a first approximation to reality, and that other more reasonable approximations should be considered in future work. We do expect that procedures developed here, based on threshold loss, will be a very good and workable first approximations indeed.

In developing our materials for this project we have been cognizant of the fact that several different kinds of technical questions would need to be resolved, and that several different kinds of audiences would need to be addressed in our final report. This has led us to seek collaboration with persons more experienced in IPI methods and the preparation of several somewhat overlapping expository papers, in addition to our technical papers. First let us consider the technical questions. The primary problem was to work with the Bayesian simultaneous decision model for the estimation of proportions in m-units, and to derive from the joint posterior distribution on these parameters, the marginal distributions for each individual element. Beyond this there was a further desire to use this posterior distribution in a full decision-theoretic analysis. Furthermore, there was a desire to utilize in the analysis, not only a general level of collateral information concerning the general level of performance of

other students in the instructional unit, but also the performance of the particular student on other instructional units. This would be particularly useful in placement testing, where a student takes short tests on several possibly highly related instructional objectives. The technical results required for these analyses are contained in papers by Novick, Lewis, and Jackson (1973, Appendix 3.1), Lewis, Wang, and Novick (1973, Appendix 3.2), Lewis, Wang, and Novick (1973, Appendix 3.3), Wang (1973, Appendix 3.4), Wang and Lewis (1973, Appendix 3.5), and Wang and Lewis (1973, Appendix 3.6), as listed in our overview summary that follows shortly. These papers tend to be rather technical in nature, and if one wishes to read more than the introductory and summary statements, one will need to follow some detailed statistical and mathematical arguments. Nonetheless, it was necessary to present this material in a rigorous technical form so that researchers wishing to extend these results would have a basis for such extension. These papers can be found in Appendix Number 3 to this report.

The general question of the application of statistical-decision theory would, in an IPI context, be one which we felt required fairly extensive discussion. While the threshold-loss function that we have adopted is certainly a useful one, and indeed we mean to have the results derived from it taken seriously, we do believe further significant improvements may be possible, using more sophisticated loss functions. To make this further work possible, we have included a rather lengthy primer on decision analysis for Individually Prescribed Instruction which comprises the content of Appendix Number 2. It is our hope that persons within IPI, and those associated with other individualized instructional programs will give this paper some study, and hopefully a dialogue will ensue among such people discussing the relative merits and demerits of various possible loss functions.

We have also, in this project, been cognizant of the fact that the procedures we are proposing are extraordinarily complicated, both theoretically and practically. Yet we intend that these procedures be adopted for classroom use by persons whose professional skill lie in instruction and not in theoretical-educational measurement or statistical decision theory. Thus we knew we would need to provide means for making these procedures available in a simple format for classroom use. We have attempted to accomplish this in two ways. First of all, in our theoretical appendix (Appendix Number 3), we have provided a set of tables (Wang, 1973, Appendix 3.4) which indeed drastically simplify the computational work in Bayesian IPI decision making. We might also refer to a paper in development by Millman (in preparation) which gives a detailed numerical example applying our methods and these tables to an IPI decision-making problem.

However, we did not feel that this approach would be entirely satisfactory. Our feeling has been, and remains, that the whole arithmetic process required for decision making will best be done, in toto, by a computer. We note that mini-computers are now in use

within IPI, and with the continuing reduction in costs of such equipment, we can speculate that in the future IPI will, in its standard form, be monitored in a computer-based environment. The question then was, could we take our decision making procedures and computerize them in such a way that this enormously complicated and sophisticated machinery could be used by persons having a verbal-theoretical understanding of what was being done, but little precise understanding of the sophisticated mathematical and statistical theory underlying the given formulas.

Fortuitously, the principal investigator has been involved, concurrently, in a project concerned with the interactive conversational analysis of data using Bayesian methods. It therefore seemed appropriate that some additional efforts be made in this area, and that the result from this additional effort be reported as part of this project. The problems faced in attempting to provide conversational language programs to monitor IPI are identical with those in other conversational statistical applications. These are discussed in detail by Novick (1973, Appendix 4.1), Isaacs (1972, Appendix 4.2), Isaacs (1973, Appendix 4.3), and Christ (1973, Appendix 4.4).

In summary then, Appendix Number 2 deals with possible further developments in decision-theory application, Appendix Number 3 deals with the core-mathematical theory underlying our proposed applications, and Appendix Number 4 deals with the computer problems involved in such applications. For most readers however, the papers of greatest interest will be those contained in Appendix Number 1. Here four papers are given which are concerned directly with the implementation of these new methods within Individually Prescribed Instruction. These papers should probably be read in the order in which they appear.

The paper by Novick and Lewis on Prescribing Test Length for Criterion-Referenced Measurement (1973, Appendix 1.1), in fact is much more general than its title would indicate. It is rather a careful laying out and consideration of all the factors which must be taken into account, both in the actual decision process and in the consideration of necessary length for criterion-referenced tests. One of the difficulties, we think, in attempting to apply decision theory in IPI, is that some of these considerations have not been discussed in the literature, and therefore there is insufficient guidance on these matters. Specifically isolated for consideration are: 1) the current level of functioning of the student, 2) the minimum acceptable criterion level to certify mastery, 3) the prior probability that a student has attained mastery, 4) the loss ratio for false positive and false negative decisions, and 5) the premium on testing time within the instructional process. Following a discussion of these topics it becomes possible to intelligently investigate the test length problem and to give some tentative recommendations. Hambleton and Novick (1973, Appendix 1.2), explicate some of the ideas contained in the paper of Novick, Lewis, and Jackson (1973, Appendix 3.1),

in a less theoretical language and provide a brief introduction to the threshold loss paradigm. Ferguson and Novick (1973, Appendix 1.3), give some details on precisely how these methods can be implemented within Individually Prescribed Instruction, and Hambleton (1973, Appendix 1.4), broadens the perspective by indicating the relevance of these methods both within IPI and for other instructional programs such as Project Plan.

II. OVERVIEW

In January of 1972, the United States Department of Health, Education, and Welfare awarded a grant of $99,492 to Dr. Melvin Novick, Director of ACT's Psychometric Research Department for a study on "New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction." The focus of the project has been on the application of certain Bayesian methods introduced by Professor D. V. Lindley of the University College London whose research has been supported in part by ACT for the past three years. Work by Melvin R. Novick, Charles Lewis, and Paul H. Jackson, on the Bayesian estimation of proportions in m groups, released as ACT Technical Bulletin No. 1, and subsequently published in Psychometrika, has been the initial take-off point for applications in this project. This work suggests that in a criterion-referenced measurement situation, an increase in precision equivalent to adding from six to twenty-five items can be attained by using the Bayesian method. In this newly defined approach, the estimation of mastery scores is replaced by the determination of the probability that the true mastery scores are larger than some specified criterion level. The result of this research is the creation of a new test theory for Individually Prescribed Instruction, and a statistical and computational technology for implementing this theory.

A bibliography of the papers completed for this project, with annotations indicating how each paper fits into the overall project development follows.

Annotated Bibliography

1.1 Novick, M. R., & Lewis, C. Prescribing test length for criterion-referenced measurement. ACT Technical Bulletin No. 18. Iowa City, Iowa: The American College Testing Program, 1973.

This bulletin demonstrates the effectiveness of the eight to twelve item criterion-referenced tests in placement pre and posttesting when the new Bayesian methods of analysis are used. It is also noted here that Bayesian techniques can be applied sequentially without modification, and thus all of the benefits of sequential analysis are available without further complicating the analysis. In the introduction, work of Millman is used to

demonstrate the inadequacy of classical analysis. Specific test length recommendations are given dependent upon (1) the loss ratio, (2) prior probabilities, and (3) the specified criterion level.

1.2 Hambleton, R. K., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10(3), 159-170.

This article describes, in nontechnical language, the ideas and methods introduced in Reference 3.1, and elaborated in Reference 3.2, to take account of the collateral information on (m-1) students to help estimate the probability that the mastery level of each m-th student is greater than the required criterion level. The central concept taken from References 3.1 and 3.2, and exposited here, is that in Individually Prescribed Instruction decisions must be based on the aposteriori probability that the student's level of functioning is greater than the prescribed criterion level. This approach is illustrated using a simple threshold loss function and a posterior marginal distribution that depends on sample, prior, and collateral information.

1.3 Ferguson, R. L., & Novick, M. R. Implementation of a Bayesian system for decision analysis in a program of Individually Prescribed Instruction. ACT Research Report No. 60. Iowa City, Iowa: The American College Testing Program, 1973.

This report provides some precise detail of how the new methods can be used in placement testing, pretesting, and posttesting. The various decision modes of IPI are identified and the precise way in which the new techniques can be implemented at each need are discussed in detail.

1.4 Hambleton, R. K. A review of testing and decision-making procedures for selected individualized instructional programs. ACT Technical Bulletin No. 15. Iowa City, Iowa: The American College Testing Program, 1973.

This bulletin discusses the similarities and differences among several approaches to individualized instruction and concludes that the new Bayesian methods will be useful in each of these approaches. Included in the survey are svstems used in Individually Prescribed Instruction, Project Plan, Mastery Learning, and approaches to computer assisted instruction.

2.1 Davis, C. E., Hickman, J., & Novick, M. R. A primer on decision analysis for Individually Prescribed Instruction. ACT Technical Bulletin No. 17. Iowa City, Iowa: The American College Testing Program, 1973.

This bulletin provides an overview of how the utilities of various outcomes can be logically combined with the aposteriori probabilities of these outcomes to provide a coherent basis for decision making. This paper illustrates in detail, results for several important utiltiy functions, thus going beyond the simple threshold loss situation. For each of these loss functions, an illustration is given of how to determine the advance-retain observed cut-score.

3.1 Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-46.

This is the fundamental theoretical paper which provides the basis in Bayesian methodology for all of the methods, theory and applications discussed in the remaining project papers. This paper was produced prior to the commencement of the project and was the basis for the project proposal.

3.2 Lewis, C., Wang, M., & Novick, M. R. Marginal distributions for the estimation of proportions in m groups. ACT Technical Bulletin No. 13. Iowa City, Iowa: The American College Testing Program, 1973.

This paper provides the key methodological development of the project -- a procedure for obtaining the aposteriori probability of mastery for each student individually from the m-group proportion method, using a well established computational approach to marginalization due to Box and Tiao.

3.3 Lewis, C., Wang, M., & Novick, M. R. A proper prior for μ_Γ in estimating proportions in m groups. ACT Technical Bulletin Supplement No. 13-1. Iowa City, Iowa: The American College Testing Program, 1973.

This supplement introduces an improvement to the basic theory of Reference 3.2 that makes it possible to incorporate prior information on the average of the group means.

3.4 Wang, M. Tables of constants for the posterior marginal estimates of proportions in m groups. ACT Technical Bulletin No. 14. Iowa City, Iowa: The American College Testing Program, 1973.

These tables make it possible to monitor Individually Prescribed Instruction without dependence on a computer.

3.5 Wang, M., & Lewis, C. Estimation of proportions in a two-way table. ACT Technical Bulletin No. 16. Iowa City, Iowa: The American College Testing Program, 1973.

This work makes it possible to take account of the collateral information contained in the test scores on t-1 objectives as well as m-1 students in estimating the proficiency of each m-th student on each t-th objective. Following transformation, a full two-way analysis of variance is used, though it is also shown in the following paper that a two-way no interaction model consistently provides almost identical results.

3.6 Wang, M., & Lewis, C. Marginal distribution for the estimation of proportions in a two-way table. ACT Technical Bulletin No. 19. Iowa City, Iowa: The American College Testing Program, 1973.

An extension of the previous paper, this bulletin provides a method of assessing the probability of a student's mastery of that objective. A no-interaction model is used here because of computational difficulties encountered in attempting marginalization with the interaction model.

4.1 Novick, M. R. High school attainment: An example of a computer-assisted Bayesian approach to data analysis. International Statistical Review, 1973, 41, 264-271.

Prepared prior to the beginning of this project, this paper demonstrates how a nonstatistician can use complex statistical techniques with the step-by-step conversational guidance of a system of Computer Assisted Data Analysis (CADA). As a result of this finding, we believe that CADA can make it possible for the classroom teacher to use the sophisticated statistical procedures developed in this project.

4.2 Isaacs, G. L. Interdialect translatability of the BASIC programming language. ACT Technical Bulletin No. 11. Iowa City, Iowa: The American College Testing Program, 1972.

A study of the BASIC programming language showing how it is possible to program in one dialect in such a way as to facilitate translation into other dialects, and thus make it possible to transport CADA programs to many different kinds of computer installations. This research makes it possible to implement CADA mode IPI programs on any adequate computer system.

4.3 Isaacs, G. L. A tabular survey of basic computer systems. ACT Technical Bulletin Supplement No. 11-1. Iowa City, Iowa: The American College Testing Program, 1973.

These tables provide a feature by feature comparison of BASIC language dialects as implemented on various computer systems with an evaluation of the adequacy of each dialect for CADA implementation. Particular emphasis is placed on chaining, string handling, and formatted output capability, as these are the BASIC features most needed in CADA. This report, completed in March, 1973, shows that a large number of computer systems are adequate for CADA applications. An on-going survey of BASIC systems indicates that many of these are being substantially improved. Most of these system updates should be completed within the next ninety days, shortly after which a revision of Technical Bulletin Supplement No. 11-1 will be prepared.

4.4 Christ, D. E. The CADA monitor. ACT Technical Bulletin No. 12. Iowa City, Iowa: The American College Testing Program, 1973.

This is a description of the Monitor used to organize the CADA package of programs. The interrelationship of the programs currently available on the Monitor is shown. Also the design philosophy, which enables the programs to be easily interconnected and used by unsophisticated investigators, is discussed. Much of the design philosophy is applicable to many other interactive situations, since its main thrust is the improvement of the man-machine interface, while minimizing programming effort.

III. SUMMARY OF RESULTS

A. The Structure of the Statistical Monitoring System and Its Implications.

The prerequisite for the introduction of a statistical monitor for IPI is a clear statement of the problem and an understanding of the evaluations that would need to be made for input into the decision-making process. In the paper entitled, "Prescribing Test Length for Criterion-Referenced Measurement" by Novick and Lewis (1973, Appendix 1.1), each of the kinds of information required for prescribing test length and making decisions based on criterion-referenced tests is discussed. The five major considerations in structuring IPI decisions are:

(1) The current level of functioning (π) of the student,
(2) The minimum advancement score (π_0) required for defining mastery of a module,
(3) Background information available on each student and on the instructional process,
(4) Relative losses incurred in making false positive and false negative decisions, and
(5) The premium on testing time within the instructional process.

In criterion-referenced testing, we think of a hypothetical (infinitely large) pool of test items relevant to a single behavioral objective. A student is considered a master of a behavioral objective if the percentage of items he would get correct over the entire pool, his <u>level of functioning</u> (π), exceeds a <u>specified criterion level</u> (π_0). Because a test contains only a small sample from that pool, errors in decision making must be expected.

As a first approximation it may be assumed that a loss a is incurred if a student is deemed a master when he is not (a false positive) loss b is incurred if he is deemed a nonmaster when he is not (a false negative), and zero loss if a correct decision is made. A coherent system of decision making is based on the aposteriori probability that the student's level of functioning is above the specified criterion level and the ratio a/b of losses associated with false positive and false negative decisions. The rule is that a student is advanced if the ratio of the probability that he is a master to the probability that he is not, exceeds the ratio of false positive to false negative losses. This is equivalent to choosing that action (advance or retain) which has the highest expected utility.

Given a prior distribution for a student's level of functioning and a given test length (sample size) it is possible to determine a <u>minimum advancement</u> (test) score required to indicate proficiency. This is the lowest score that will yield an aposteriori probability of

mastery large enough to justify an advancement decision. The details of this analysis are given in this paper (Novick and Lewis, 1973; Appendix 1.1). Consideration is given to specified criterion levels of .70, .75, .80, and .85. The loss ratios are assumed to take the values 1.5, 2.0, 2.5, or 3.0. Thus it is assumed that typically a loss incurred for a false positive decision will be at least one and one-half times that for a false negative decision. Various prior distributions are considered for each specified criterion level. Generally it is assumed that the prior distribution will have a mean value near the specified criterion level and typically and desirably slightly larger. For each analysis with a particular expected value for the prior distribution, four different priors are considered with varying degrees of certainty in the prior distribution. The results of these analyses are summarized in a set of tables, seven through eleven (Appendix 1.1, Pages 20, 23, 26 and 28). For each combination of specified criterion level, loss ratio, and prior distribution, a recommended test length and minimum advancement score are given. At the end of each table, some general recommendations are given which seem to be reasonable for a wide range of prior distributions for the particular specified criterion levels and loss ratios. For example, with a π_0 value of .70, and a prior distribution having expectation of .70, the general recommendation for a loss ratio of 1.5 is a test of eight items with the requirement of six out of eight correct for advancement. For a loss ratio of 2.0, a test of 13 items is recommended with a score of ten being required for advancement. The ratio of 2.5 requires a test of 14 items with eleven correct for advancement, and a loss ratio of 3.0 requires a test of 15 items with 12 correct for advancement.

We suspect that these particular recommendations will be pleasant news for IPI people. If the loss ratio is as small as 1.5, as it may well be in some situations, the indication here is that an 8-item test will be satisfactory. Even with a loss ratio of 2.0, a 13-item test will do. Loss ratios of 2.5 and 3.0 do not call for greatly increased test lengths, however loss ratios as high as this may indicate that the structure of the unit and its relationship to other units could profitably be reevaluated. We shall discuss this and similar questions later in this report. With a π_0 value of .75, and a prior distribution with expectation .75, and a loss ratio of 1.5, the recommended test length is 10 items with a minimum advancement score of 8. While this will likely be thought of as a very acceptable test length, the situation does not remain as favorable when the loss ratio rises to 2.0. Here the test length recommendation is for a test of 25 items with a minimum advancement score of 16. As is indicated later in the paper, this situation can be improved by training the group to a higher average level of performance. For a π_0 value of .80 and a prior expectation of .80, reasonably satisfactory test length specifications are obtainable provided the loss ratio does not exceed 2.0. Specifically, for a loss ratio of 1.5, a seven-item test is deemed adequate with a minimum advancement score of six. The loss ratio of 2.0 on an eight-item test will be adequate with a minimum advancement

score of seven. With loss ratios 2.5 and 3.0, tests of length 20 and 22, with minimum advancement scores of 17 and 19, respectively, are recommended.

When the situation does suggest the appropriateness of high loss ratios, a considerable reduction in the required test length can be obtained by seeing to it that the means of the prior distribution is rather higher than the specified criterion level. For example, with a π_0 value of .80 and a prior expectation for π of .85, 12 and 13-item tests will be adequate for loss ratios of 2.5 and 3.0. This compares with test lengths 20 and 22 where the prior expectation of π was .80. The Novick-Lewis paper contains detailed discussions of the relationship of the various input variables beyond that which we shall not cover here, though we urge readers to study that paper carefully. We shall consider here only some of the broad implications of this study for the very structure of Individually Prescribed Instruction. First we would note that there has been a definite tendency in IPI to require relatively high criterion levels; typically, the value .85 is used. One might well speculate whether this really reflects a perceived need for a high criterion level, or whether it is, in fact, a function of a high loss ratio combined with a desire for a short test length. Only when we get to the point that required loss ratios and criterion levels can be independently evaluated will it be possible to use tables such as the ones presented in this paper. We might speculate, and indeed hope, that for most situations a specified criterion level of .75 will be adequate, but that a loss ratio greater than one, possibly 1.5 or 2.0, will be appropriate.

The tables presented in this paper also have a great implication for the amount of time that might best be put into an individual training unit. When loss ratios are high, it may well be highly advantageous to strengthen the training program to the extent that the mean output is well above the specified criterion level. This will make it possible to use short tests, or alternatively, will generally reduce the risk of incorrect classification. This will of course be more expensive and this investment must be balanced out against the reduction of cost of testing and the reduction in the expected loss due to incorrect decision.

Another implication of these tables is that the training module should also be structured so that very high loss ratios are <u>not</u> appropriate. This will be accomplished by seeing to it that individual modules are not overly dependent on preceding ones. Here again, one is balancing off changes in the module itself against changes in the criterion-referenced testing.

We would emphasize again that the primary purpose of this paper is to provide a structure for an intelligent discussion of decision making within IPI, including the question of prescribing test length. The results contained here,.we think will be useful, but they should in no way be considered to be definitive. We do not know what loss

functions or specified criterion levels are appropriate. We do know that the loss function adopted here is only a first approximation to reality. We are reticent to consider more complicated loss functions at this point, because of the difficulty of getting reliable judgments as to which of these may be appropriate. Our hope is that this paper will initiate some substantial discussions among IPI people, so that further papers of this kind can be attempted, based upon more definitive evaluations of loss ratios, specified criterion levels, prior distributions, and loss functions. Some of these points are touched upon in the summary of the Novick-Lewis paper.

The paper entitled "Toward an Integration of Theory and Method for Criterion-Referenced Tests" by Hambleton and Novick (1973, Appendix 1.2), provides a first attempt at tying previous discussions of criterion-referenced testing and Individually Prescribed Instruction to formal decision theory. Generally definitions given by Glaser and Nitko (1971) are taken as starting points for this marriage of theory and practice. According to Glaser and Nitko, "a criterion-referenced test is one that is deliberately constructed so as to yield the measurements that are directly interpretable in terms of specific performance standards." The performance standards are usually specified by defining some domain of tasks that the students should perform. Representative samples of tasks from this domain are organized into a test. Measurements are taken and used to make a statement about the performance of each individual, relative to that domain. Thus the quantity of interest here is the level of functioning of the individual student, which is defined as the proportion of items that he would answer correctly in a hypothetical infinitely large population of items relevant to the specific behavioral objective which this test examines. It is then further assumed that as a first approximation, it is appropriate to specify some value on this scale, 0 to 1, defined by this level of functioning, and to call that point the minimum criterion level. Persons with levels of functioning above this level, are considered to be masters and those below it are considered to be nonmasters. The decision process is thought to be one of deciding in which of these two categories an individual person belongs. It is recognized that this arbitrary dichotomization is somewhat unreal, however under certain circumstances it will represent quite a reasonable first approximation to reality.

The third paper in Appendix Number 1 is entitled "Implementation of a Bayesian System for Decision Analysis in a Program of Individually Prescribed Instruction" by Ferguson and Novick (1973, Appendix 1.3). This paper should be particularly useful to those persons who have some knowledge of IPI, but require some further details in order to understand how our new statistical procedures would be implemented. The paper begins with a brief description of the structure of IPI, with a detailed description of the IPI mathematics program. There then follows a specific description of how the instructional decision-making process would function. It is pointed out that in IPI there is a great deal of information available about the instructional program. Quite specific information is available concerning the

distribution of the percentage of items answered correctly by students, and it is thus possible to make inferences about the true level of functioning of each student and the mean and standard deviation of these true values in the population of students. It is further pointed out that if the instructional program were completely efficient and the students were without human frailty, there would be no variation in true level of functioning of students on posttests. A student would remain in a unit only until that instant at which his level of functioning attains a prespecified criterion. However nothing approaching this is possible with present instruction technology. In the real world of IPI there will be some variation of true levels of functioning among students on posttests. And it is pointed out that this background information can be combined with direct observational information to improve the decision-making process. It is also noted that other background information can be used, namely that involving the performance of the student on tests of other skills. It is argued that surely a person scoring highly on t-1 subtests, and a little less highly on t-th would, we suspect, have a true score on the t-th test higher than his observed score, and that this somewhat lower observed score might be due in part to bad luck or carelessness. It is pointed out that the method for performing these analyses is available from the work of Wang and Lewis (1973; Appendix 3.5 and 3.6).

The report ends by illustrating the kind of format that can be used to transmit information to the instructional manager. Currently following an IPI posttest, the manager receives a skill profile on each student. On this profile the percentage correct that the student got on each of the skills is reported. And from this the instructional manager decides which skills the student must redo. Under the proposed change the posttest profile would not consist of these percentage correct scores, but rather, for each skill, the probability that the student's true level of functioning is greater than the specified criterion level. Thus while the information to be fed to the instructional manager is somewhat different than it has been in the past, it is certainly no more complex. It will be necessary to teach instructional managers what these new numbers are, what they mean, and how they are to be used. But this process should not be terribly difficult. Indeed there may well be enough instructional material in this particular report to accomplish this task.

The fourth paper appearing in Appendix Number 1 is "A Review of Testing and Decision-Making Procedures for Selected Individualized Instructional Programs" by Hambleton (1973). While the current research effort has been directed primarily at one particular individualized program, namely IPI, our view is that these same methods can be used in other programs, namely Project PLAN and the Mastery Learning Program, as well as in various approaches to computer-assisted instruction. The survey by Hambleton gives us enough of a picture of each of these programs to confirm this belief, and furthermore suggests that none of these programs currently has any sort of well-developed decision process. Undoubtedly a major undertaking would be required to implement the Bayesian decision-theoretic system in each of these. There is no question in our mind but that this would be useful.

B. Statistical Decision Theory for Individually Prescribed Instruction

The papers in Section A have proposed a statistical monitoring system for IPI in the framework of statistical decision theory. In particular, a threshold loss structure for the problem was proposed and the implications of that loss structure were investigated. At the same time it was pointed out that this is only one of many possible loss structures for this problem and that indeed it should only be considered as a first approximation to an appropriate loss structure.

In Appendix 2 we present a paper entitled "A Primer on Decision Analysis for IPI". The purpose of this paper is to provide a brief semi-technical presentation of decision theory set within the context of IPI, so that persons concerned with IPI can learn enough of decision theory to understand how such methods can be applied. The paper begins with a formal statement of decision theory in a two decision situation and demonstrates the application of normal and extensive form analysis. Computations for both types of analysis are presented in complete detail so that the reader can see how each operates. It is then pointed out that under general conditions, both normal and extensive form of analysis will always lead to the same decision, and since extensive form analysis is the easier to do, it therefore becomes possible and desirable to adopt it as the standard procedure.

In extensive form analysis the first task is to compute the posterior probability distribution of the unknown parameter given the prior distribution and the data. Once this is done this probability distribution can be combined with the statement of the loss structure to arrive at an extensive form decision. The discussion then turns to the use of extensive form analysis with continuous posterior distributions and specific applications and examples relevant to IPI are given. This continuous form for the prior distribution is the one that has been discussed previously in the expository papers. Particular emphasis is placed in the discussion here on procedures for determining cutting scores, that is the point in the observed score continuum above which a student should be deemed a master and below which he should be deemed a nonmaster. This complements the work in Appendix 1.1.

There then follows a reasonably complete discussion of utility theory which indeed makes a generalization of threshold loss possible. In this way it is possible to have different utility for true positive and true negative decisions. The paper then discusses a linear utility function which may be useful in some situations followed by a discussion of quadratic utility and exponential utility.

The most reasonable appearing utility function and one which may indeed be most appropriate for IPI application is the squared

exponential utility function. A brief description of this utility function is given. There then follows a discussion of a three action problem in which the possible decisions are

(1) The student should be sent back one module,
(2) He should be retained in the present module,
(3) He should be advanced to the next module.

A complete analysis for this situation is discussed within the framework of threshold utility. It is then discussed in the framework of linear utility.

The final section of the report deals briefly with the question of deciding whether or not it is useful, at a particular point in a sequential testing environment, to take another observation, that is have the student answer another item.

It is a general property of Bayesian inference that the analysis is identical whether observations are taken all at once or taken sequentially with the possibility of stopping whenever a decision can legitimately be made. For example, if the appropriate rule is that a student can be advanced if he gets seven out of eight and the items are being administered sequentially, he could terminate test taking if he were to answer the first seven items correctly or if, at any time, his total of incorrect responses exceeded one. In a more general framework when exact Bayesian solutions are used rather than the approximate ones discussed here and in Appendix 1.1, the formal procedure for deciding whether or not an additional item should be administered involves comparing the expected value of information to be obtained from that item with the cost of administering this item. This, of course, is that the value of information and the cost of administering an item have been put on a common scale. In practice this is a very difficult thing to do. We have not in this report, therefore, considered costs involved in testing but in theory this could be done.

C. Theoretical Developments

The thrust of the theoretical development for this project is based on a specialization to the "Estimation of Proportions in m Groups" by Novick, Lewis and Jackson (1973, Appendix 3.1), of the methods of simultaneous estimation developed in a Bayesian context by Lindley (1971) and in other contexts by Stein (1962), Robbins (1955), and others. The Bayesian model employed here begins with the assumption that each student is administered a random sample of items from a population of items and that the scoring is binary right/wrong. It is assumed throughout that all students received the same number of items. In order to greatly simplify the analysis, the random variable X, which is the number of correct responses of a student, is transformed to a new variable, G. The transformation employed is the well-known root arcsine transformation which in its simplest form is the arcsine of the square root of the ratio x/n. This transformation in a more complex form due to Freeman and Tukey (1950) has the advantage of providing a random variable which for even moderate values of n (eight will generally do for our purposes) is such as to have a known variance as a function of n alone. Thus the data then consists of m observations from m different normal popualtions (persons) with known variances, but unknown means γ_i. The mean value, γ_i in each of these populations is the corresponding arcsine transformation of the level of functioning π_i for the individual student. The problem then is to simultaneously estimate the m values γ_i .

If it can be assumed that no prior information exists which can differentiate one student from another, then our joint prior distribution on the set of parameters $\underset{\sim}{\gamma}$ will be exchangeable and this, mathematically, will be equivalent to the assumption that these students were randomly sampled from some population. If we make this assumption and further strengthen the model by assuming that the population is normal with mean μ_Γ and variance ϕ_Γ the model is complete.

The Novick, Lewis, and Jackson paper (1973) shows how it is possible to introduce prior information on the variance ϕ_Γ of this distribution and the later note (Appendix 3.3) shows how it is possible to incorporate prior information on μ_Γ .

When this is done and the data are put into Bayes theorem, the result is a joint posterior distribution on the set of ability parameters γ. In the Novick, Lewis and Jackson paper (1973), the strategy at this point was to obtain a joint model estimate of the $\underset{\sim}{\gamma}$ whose elements are the γ_i and to then transform these elements into estimates for the individual values π_i . This solution is consistent with what is usually required in norm-referenced testing but is inadequate for the criterion-referenced requirements of Individually Prescribed Instruction.

The central technical development of this project is contained in Appendix 3.2. In this paper, the work of Novick, Lewis and Jackson (1973) is extended by taking the posterior joint distribution of the γ_i and obtaining from it the marginal distribution of each γ_i . The approach here is entirely numerical, as it is not possible to obtain a closed form expression for these marginal distributions. After several unsuccessful attempts, an approach used originally by Box and Tiao (1968) and by Hill (1956) was found to be adequate. This approach involved obtaining the marginal distribution of γ_i conditional on ϕ_Γ, and the marginal distribution of ϕ_Γ alone, and then obtaining marginal distribution of γ_i (unconditional) by numerically integrating with respect to ϕ_Γ . A computer program of some complexity was written to accomplish this numerical integration.

If we assume that IPI is being monitored in a computerized environment, this computer program can be incorporated as a subroutine and its complexity becomes a matter of no concern since the user need have no direct contact with it. On the other hand there is often a desire to monitor IPI offline, in which case it would not be possible to make the necessary computations of the marginal distributions. In order to make this possible an asymptotic expression for the marginal distribution of γ_i was obtained. It turned out that γ_i can be well approximated by a normal distribution provided the number of items is eight or more, and π_i is not too near zero or one. Unfortunately the mean and variance of this normal distribution cannot be obtained as closed-form expressions and must in fact be calculated numerically. It turned out, again, that these computations are complex. Therefore in order to make it possible to do offline IPI monitoring, it was necessary to construct a set of tables for the mean and second raw moment of this asymptotic distribution. With these tables (Wang, 1973; Appendix 3.3) there is little difficulty in performing the necessary calculations.

It may be appropriate at this point to make some remarks concerning the force of the exchangeability assumption. This assumption requires that when the analysis is performed, we have no differential information about students. This assumption may not be valid for some students. It may first be violated in that we have a record of prior performance on each student and it may be the case that some students are typically repeaters, in that they typically must take posttests two or three times, while others may typically get a pass on a posttest the first time through. If this is the case, it may well be useful to categorize students in this way and to do separate analyses for these subgroups of students. However, where no consistent pattern can be found, it should be satisfactory to do the analysis on all students at one time.

A second situation in which this assumption may be invalid is for students who have in fact already failed this examination on one occasion, and are retaking it as repeaters. If it is known that repeaters differ from first time test takers in any particular way,

it may, in fact, be necessary to treat repeaters as a separate group. These considerations can of course only be determined by studying data from a particular IPI application. Nevertheless, it is important that this assumption be well understood by those attempting application of this technique.

Appendix Number 3.5 and 3.6 by Wang and Lewis (1973 a, b), represents a significant advance in statistical technology for IPI monitoring. While the theory contained in these papers is complete, it has not been subject to even the modest empirical study that the simpler theory has been, and as a result we do not have as clear an indication at this point as to what further practical improvements on the simpler theory are made possible here. The idea behind the work in these references is that it is possible to gain information about a particular student's ability on a particular behavioral objective not only from the fact that he was in an instructional program with a group of other students, all of whom received the same training and that therefore they can be expected to be at a roughly similar level, but also it is possible to note that this student has been trained and is now being tested on other behavioral objectives at the same time, and that his ability on the t-th behavioral objective will surely be reflected to some extent by his performance on the other t-1 objectives. This will be true if there is any relationship at all between the t objectives, as there typically is. Thus the whole Kelley (1923) approach can be used to adopt collateral information from these other objectives for estimation of the t-th objective on each student.

The mathematical approach here, following root arcsine transformation, is to utilize a full two-way analysis of variance model, which in the first instance was studied with possible interaction. The first Wang and Lewis paper (1973 a) shows how to analyze this model, and shows how to get out point estimates. The work here is in a way similar to some previous work done by Lindley though the specialization here yields much more simple results due to the fact that the variances known.

In these papers, a two-way no-interaction model is also used and it is found that on the data sets investigated almost no difference in point estimates were obtained from the interaction and no interaction models. A possibly wild speculation here is that the arcsine transformation not only gives homogeneous variance, and to some extent normality of distribution, but also has the desirable side effect of tending to yield additivity.

In the second Wang and Lewis paper (1973 b), the no-interaction model is employed in order to get out marginal distributions for the γ_i in a manner very similar to that used in Appendix 3.2. These posterior marginal distributions of the γ_i can then be used in a decision analysis in precisely the same way as posterior distributions from simpler analyses. It proved impossible to obtain marginal

distributions numerically with the interaction model and so the no-interaction model was used.

The work in Appendix Number 3.5 and 3.6 should be useful both in posttest and placement applications where inference needs to be made simultaneously on several behavioral objectives for each student. From the work in Appendix 3.1 we have a reasonably good idea of the benefit to be gained from the collateral information on other students taking the test. However, we have not as yet been able to give sufficient thought to have any idea of how much information is gained and how much resultant decrease in testing can be accomplished by using information on the t-1 other behavioral objectives for each student. In any event there is some limitation here, in that in order to use the arcsine transformation theory with a normality assumption, a sample of eight items seems to be almost necessary if we are to work at all away from the center of the distribution. Thus even with a π_0 value of .75, an n of 8 seems desirable, if not absolutely necessary, while with a π_0 value of .85, an item sample of size 8 seems at best to be barely adequate to justify the assumptions of the model.

At this point it seems clear that the theory will be extremely useful, but that it will be necessary to study and determine the kinds of distributions that are to be found in practice and to determine as a result of this what test lengths and decision rules will be appropriate. The discussion in Appendix 1.1 barely begins to tap the question of test length specification. However, it will be exceedingly difficult to do any further work on this without a close look at real data.

D. Computer Techniques

Both the complexity of the solutions provided in (C) and the desirability of maintaining continuously updated records on each student, virtually necessitates the implementation of the statistical monitoring system for IPI within a computer environment. We therefore wish to discuss certain issues that arise once this position is taken. First we wish to indicate that we feel that a computerized application is feasible without the existence of highly technical computer personnel at each installation.

What we have in mind is that it should be possible to centrally develop fully conversational interactive programs for monitoring IPI which are such that the instructional manager will need to know almost nothing about the mechanics of the monitoring system beyond that contained in the references in Appendix Number 1.

The possibility of using a computer to lead a researcher or manager through a complex statistical analysis on a step-by-step basis as in computer-assisted instruction has been demonstrated by Novick (1973). His system of Computer Assisted Data Analysis (CADA) (Appendix 4.1), has been used to make available some very highly complex Bayesian statistical methods, and these methods have been demonstrated as being usable by relatively unsophisticated users. Appendix 4.1, developed outside this project, shows one application of CADA, and we believe it is such as to demonstrate that the typical IPI instructional manager could do the kind of complex work we are discussing here with very little training. In order to monitor his IPI class, he would need only to respond to questions put to him by the computer and to use the computer output in much the way he has used other information in the past. This is described briefly in Appendix 1.3.

Once this position is accepted, it then becomes a question of preparing IPI management programs for use on a variety of computer hardware. This introduces a major problem, in that various computers have rather different interactive capabilities. Two possibilities were examined. One involved the use of interactive FORTRAN and the second, the use of the BASIC programming language. Interactive FORTRAN has the advantage of being somewhat more highly standardized from one computer to another, whereas BASIC seems to have rather different dialects for each hardware system. The disadvantage of interactive FORTRAN is that it operates only on large scale computers and not on mini-computers, and furthermore it is not, in practice, available on very many computers at present. On the other hand, every mini-computer manufacturer has one or more versions of BASIC available for their machines. Therefore, a close look was made at the BASIC programming language and its various dialects, to see if it would be possible to write in one dialect of BASIC and to reprogram, or translate, into other dialects. A survey undertaken by Isaacs (1972; Appendix 4.2) came to the conclusion that it would be possible to select one dialect of BASIC as the core dialect, to ignore some of

the sophisticated features of that dialect, and to then write programs which could be very easily translated into any other dialect. At present we have been working with the Hewlett Packard HP2000C dialect of BASIC, but we could equally well work with the Digital Equipment Corporation PDP11/40 dialect. In either case, if care is taken in the writing of the program, translation to the other dialect or indeed to any other dialects would be very easy.

The second report written by Isaacs (1973, Appendix 4.3), carefully surveys all of the BASIC dialects available in March of 1973, and indicates the strength and weakness of each of these. The conclusion is that most dialects have sufficient capability for CADA application, and that therefore IPI monitoring could be accomplished with them. It should be noted that BASIC dialects for mini-computers are under constant states of revision and that by the time this report is filed, many of the surveyed dialects will be much stronger than they were in March of 1973. In particular, we would note that the Wang 2000 super desk calculator would seem to come very close to having the capability for CADA application. If this is true, then the cost of the computerized IPI management becomes almost trivial.

The final paper in this report is a technical description of the CADA monitor indicating how subroutines can be chained to the monitor in BASIC and how it is possible to continually update and improve the monitor without disturbing the system in operation.

The statistical methodology developed here, the availability of relatively inexpensive computational machinery, and the clear understanding and explication of a theoretical structure for IPI, we think now makes it clearly possible and highly desirable to introduce a structured management for IPI. The theory presented here is still just that, theory. To make it work it is now necessary to implement these decision-making procedures within an ongoing IPI operation. No doubt such application will result in the refinement of the theory and hopefully in its improvement.

REFERENCES

Box, G.E.P., & Tiao, F. C. Bayesian estimation of means for the random effects model. Journal of the American Statistical Association, 1968, 63, 174-181.

Christ, D. E. The CADA monitor. ACT Technical Bulletin No. 12. Iowa City, Iowa: The American College Testing Program, 1973.

Davis, C. E., Hickman, J., & Novick, M. R. A primer on decision analysis for Individually Prescribed Instruction. ACT Technical Bulletin No. 17. Iowa City, Iowa: The American College Testing Program, 1973.

Ferguson, R. L., & Novick, M. R. Implementation of a Bayesian system for decision analysis in a program of Individually Prescribed Instruction. ACT Research Report No. 60. Iowa City, Iowa: The American College Testing Program, 1973.

Freeman, M. F., & Tukey, J. W. Transformations related to the angular and the square root. The Annals of Mathematical Statistics, 1950, 21, 607-611.

Hambleton, R. K. A review of testing and decision-making procedures for selected individualized instructional programs. ACT Technical Bulletin No. 15. Iowa City, Iowa: The American College Testing Program, 1973.

Hambleton, R. K., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10(3), 159-170.

Hill, B. M. Inference about variance components in the one-way model. Journal of the American Statistical Association, 1956, 60, 806-825.

Isaacs, G. L. Interdialect translatability of the BASIC programming language. ACT Technical Bulletin No. 11. Iowa City, Iowa: The American College Testing Program, 1972.

Isaacs, G. L. A tabular survey of basic computer systems. ACT Technical Bulletin Supplement No. 11-1. Iowa City, Iowa: The American College Testing Program, 1973.

Kelley, T. L. Statistical Methods. New York: The Macmillan Company, 1923.

Lewis, C., Wang, M., & Novick, M. R. Marginal distributions for the estimation of proportions in m groups. ACT Technical Bulletin No. 13. Iowa City, Iowa: The American College Testing Program, 1973.

Lewis, C., Wang, M., & Novick, M. R. A proper prior for μ_Γ in estimating proportions in m groups. ACT Technical Bulletin Supplement No. 13-1. Iowa City, Iowa: The American College Testing Program, 1973.

Lindley, D. V. The estimation of many parameters. Foundations of statistical inference. (V. P. Godambe and D. A. Sprott, Eds.) Toronto: Holt, Rinehart, and Winston, 1971, 435-455.

Lindley, D. V. Discussion following Stein, C. M. Confidence sets for the mean of a multivariate normal distribution. Journal of the Royal Statistical Society, (Series B), 1962, 24, 265-296.

Lindley, D. V., & Smith, A.F.M. Bayesian estimates for the linear model. The Journal of Royal Statistical Society, (Series B), 1972, 34, 1-41.

Millman, J. Criterion-referenced measurement. (In preparation).

Novick, M. R. Multiparameter Bayesian indifference procedure. Journal of the Royal Statistical Society, (Series B), 1962, 24, 265-296.

Novick, M. R.. High school attainment: An example of a computer-assisted Bayesian approach to data analysis. International Statistical Review, 1973, 41, 264-271.

Novick, M. R., & Lewis, C. Prescribing test length for criterion-referenced measurement. ACT Technical Bulletin No. 18. Iowa City, Iowa: The American College Testing Program, 1973.

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-46.

Robbins, H. An empirical Bayes approach to statistics. Proceedings of 3rd Berkeley Symposium on Mathematics, Statistics, and Probability, 1, 157, University of California Press, 1955.

Stein, C. M. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proceedings of 3rd Berkeley Symposium, 1956, 1, 197-206.

Stein, C. M. Confidence sets for the mean of a multivariate normal distribution. Journal of the Royal Statistical Society, (Series B), 1962, 24, 265-296.

Wang, M. Tables of constants for the posterior marginal estimates of proportions in m groups. ACT Technical Bulletin No. 14. Iowa City, Iowa: The American College Testing Program, 1973.

Wang, M., & Lewis, C. Estimation of proportions in a two-way table. ACT Technical Bulletin No. 16. Iowa City, Iowa: The American College Testing Program, 1973.

Wang, M., & Lewis, C. Marginal distribution for the estimation of proportions in a two-way table. ACT Technical Bulletin No. 19. Iowa City, Iowa: The American College Testing Program, 1973.

Appendix Number 1

to

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

Appendix Number 1

to

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

ACT TECHNICAL BULLETIN NO. 18

PRESCRIBING TEST LENGTH FOR CRITERION-REFERENCED MEASUREMENT

by

Melvin R. Novick and Charles Lewis

The American College Testing Program
and
The University of Iowa

The University of Illinois

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

January, 1974

Prescribing Test Length For Criterion-Referenced Measurement

I. Posttests

by

Melvin R. Novick and Charles Lewis

The American College Testing Program
and
The University of Iowa

The University of Illinois

Introduction

In a program of Individually Prescribed Instruction (IPI), where a student's progress through each level of a program of study is governed by his performance on a test dealing with individual behavioral objectives, there is considerable value in keeping the number of items on each test at a minimum. The specified test length for each objective must, however, be adequate to provide sufficient information regarding the student's degree of mastery of the behavioral objective being tested. Just what the minimum acceptable length will be depends on the manner in which test information is used to make decisions about individual students, the level of functioning required for defining mastery of an objective, the relative losses incurred in making false positive and false negative decisions, the background information available on the student and on the instructional process, and the premium on testing time within the instructional process. Our purpose in

---

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy. We are grateful to Charles Davis and Nancy Petersen for helpful comments and computations. This paper will be published in the CSE Monograph Series in Evaluation, Number 3, a publication supported in part by the National Institute of Education and by the American Educational Research Association.

this paper is to discuss these issues and provide some broad guidelines for test-length specification for IPI posttests. These specifications will be tentative because of unresolved substantive and methodological issues, but we believe that they should provide some improvement on current practice. A separate, and rather more complex treatment will be required for placement and pretest length specification.

Background

In a criterion-referenced measurement approach to Individually Prescribed Instruction, we imagine a population of test items, having mixed item difficulty, dealing with a particular objective and an ideal decision which advances a student past this objective if he is able to answer at least a given percentage of the items in the population. This minimum passing percentage, the so-called <u>criterion level</u>, simply reflects the degree of mastery deemed sufficient for this objective (although it implicitly involves the difficulty of the items as well). The actual percentage of items that a person would answer correctly in the population of items is called his <u>level of functioning</u>. In practice, the advancement-retention decision must be made from a small sample of observations (test items), and, hence, errors in the decision process must be expected.

One common treatment of the test length problem in a criterion-referenced measurement context has been given by Millman (1972). He studied a standard decision rule which advances the student if the percent of items correctly answered on a test equals or exceeds the required criterion level. Here it is assumed that the items on the test may be treated as a random sample from the population of interest, so that the obtained percentage correct is a useful estimate of the true population percentage for the student. Using binomial probability

Table 1

Percent of Students Expected To Be Incorrectly

Advanced or Retained

Specifieu Criterion Level .70

Advancement Score	No. of Test Items	Student's True Level of Functioning* 50	55	60	65	70	75	80	85	90	95
6	7	6	10	16	23	67	55	42	28	15	4
6	8	15	22	32	43	45	32	20	11	4	1
7	9	9	15	23	34	54	40	26	14	5	1
7	10	17	27	38	51	35	22	12	5	1	-
8	11	11	19	30	43	43	29	16	7	2	-
9	12	7	13	23	35	51	35	20	9	3	-
10	13	5	9	17	28	58	42	25	12	3	-
11	14	3	6	12	22	64	48	30	15	4	-
12	15	2	4	9	17	70	54	35	18	6	-

Specified Criterion Level .75

Advancement Score	No. of Test Items	Student's True Level of Functioning* 50	55	60	65	70	75	80	85	90	95
6	8	15	22	32	43	55	32	20	11	4	1
7	9	9	15	23	34	46	40	26	14	5	1
8	10	6	10	17	26	38	47	32	18	7	1
9	11	3	7	12	20	31	55	38	22	9	2
9	12	7	13	23	35	49	35	20	9	3	-
16	20	1	2	5	12	24	58	37	17	4	-
17	21	-	1	4	9	20	63	41	20	5	-
18	22	-	1	3	7	17	68	46	23	6	-

Table 1 (continued)

Specified Criterion Level .80

Advancement Score	No. of Test Items	50	55	60	65	70	75	80	85	90	95
		Student's True Level of Functioning*									
6	7	6	10	16	23	33	45	42	28	15	4
7	8	4	7	11	17	26	37	50	34	19	6
8	9	2	4	7	12	20	30	56	40	23	7
8	10	6	10	17	26	38	53	32	18	7	1
9	11	3	7	12	20	31	46	38	22	9	2
10	12	2	4	8	15	25	39	44	26	11	2
11	13	1	3	6	11	20	33	50	31	13	2
12	15	2	4	9	17	30	46	35	18	6	-
17	20	-	1	2	4	11	23	59	35	13	2
19	22	-	-	1	3	7	16	67	42	17	2

Specified Criterion Level .85

Advancement Score	No. of Test Items	50	55	60	65	70	75	80	85	90	95
		Student's True Level of Functioning*									
7	8	4	7	11	17	26	37	50	34	19	6
8	9	2	4	7	12	20	30	44	40	23	7
9	10	1	2	5	9	15	24	38	46	26	9
10	11	1	1	3	6	11	20	32	51	30	10
11	12	-	1	2	4	9	16	28	56	34	12
17	19	-	-	1	2	5	11	24	56	29	7
19	21	-	-	-	1	3	8	18	63	35	8

*The true level of functioning is the percent of items a student would be able to answer correctly if he were given the entire universe of items.

Students having true level of functioning values less than the specified criterion level should fail a test composed of all items from this universe. However, on any given test of finite length, some of these students will get more than the minimum advancement percent of the items correct and be considered as "passers". The expected percent of such incorrect advancements are given in the body of the table to the left of the dotted line.

Students having true level of functioning values equal to or greater than the minimum advancement percent should pass such a test. The percent of these students who will be incorrectly retained are shown in the table to the right of the dotted line.

tables, Millman obtained the probability that a student with a given <u>true</u> level of functioning would be incorrectly advanced or retained by this procedure.

Table 1 expands on some of Millman's computations and gives the conditional probability of incorrect advancement or retention for a variety of true levels, test lengths, and minimum passing percentages. The first impression this table provides is that a substantial proportion (sometimes more than half) of the students with true levels close to, or at the criterion level, will be incorrectly advanced or retained, at least for the test lengths considered. There appears to be a slight improvement in accuracy of decision as the test length increases from 8 to 22 items, although this effect is largely hidden by fluctuation in the probabilities, due to changes in the percentage correct required for advancement. For example, with a criterion level of .7, the percentage correct required for advancement is .75, .78, .70, .73, or .75 for test lengths of 8, 9, 10, 11, or 12 items, respectively. This brings up a question as to the optimality of the decision procedure assumed in Table 1. To provide a framework for answering this question, let us consider some of the issues involved.

Suppose seven out of eight were taken as the <u>minimum advancement score</u> when the criterion level os .75; the probability of incorrect advancement would decrease substantially for all students with true levels below the criterion level. This is shown in Table 2. On the other hand, those above .75 suffer a substantial increase in their chances of being incorrectly retained. Apparently, a more general framework is required before even the decision procedure can be chosen, much less any judgment made concerning minimum test length. This framework would need to take into account on which side of .75 small expected errors were considered to be more important.

Table 2

Percent of Students Expected To Be Incorrectly

Advanced or Retained

Criterion Level = .75 Test Length = 8

Advancement Score	True Level 50	55	60	65	70	75	80	85	90	95
6	15	22	32	43	55	32	20	11	4	1
7	4	7	11	11	26	63	50	34	19	6

A Framework For Specifying Test Length

Table 1 is very helpful in identifying the seriousness of the problem of short tests. From a practical point of view, however, a solution to the problem must involve looking at a different conditional probability, and abandoning the simple decision procedure that Millman has so convincingly demonstrated to be inadequate. Instead of the probability that a student will attain a particular test score, given his true level, it is the probability that a student's true level of functioning exceeds the specified criterion level, given his test score, which is required in making a decision. In other words, it is the test score--not the true level--which is given (i.e. observed), and which is the basis for any decision to advance or retain the student. Thus, a student should be advanced only if the probability that he has attained or surpassed the criterion level, <u>given his test score</u>, is sufficiently high. To obtain the necessary probability, an application of Bayes theorem is required. In such an analysis, prior knowledge (expressed in probabilistic terms) of the student's true level of functioning is combined with the (binomial) model information relating the observed test score to true level; and, the result is a posterior probability

distribution for true level of functioning, given test score. The probability this distribution assigns to levels above the criterion is the quantity of interest. In this formulation, the problem can be described as selecting a minimum sample size and an advancement score, so that students attaining that score will then have a sufficiently high probability of having at least the minimum required level of functioning.

As a first approximation, let us suppose our knowledge of a student's true level of functioning is vague, prior to having his test results. If this state of knowledge is characterized by selecting a uniform distribution on the interval from zero to unity for true level, π, Bayes theorem provides the posterior probabilities listed in Table 3 for various scores and test lengths. The posterior distributions on which these probabilities are based all belong to the Beta family, and the parameters in each case are those given in the table, primarily for future reference.

To generate a decision procedure on the basis of Table 3, we must select a criterion level (π_o) and a minimum acceptable probability that a student's true level (π) exceeds this criterion. Thus, for example, we might take $\pi_o = .80$ and the minimum acceptable $\text{Prob}(\pi \geq \pi_o | x, n) = .50$, where x is test score and n is test length. We would then be saying that we wanted to advance the student only if we were at least 50% sure that his level of functioning was above .80. Then, using Table 3, we see that with $n = 8$, all students having $x \geq 7$ would advance to the next objective, but not those with $x = 6$. For a test of 12 items, the minimum advancement score would be 10 correct.

Note, however, that if we required 80% assurance that the true level of functioning was above .80, [$\text{Prob}(\pi \geq .80) \geq .80$], then even those with eleven correct responses to twelve items would not be advanced. We think

Table 3

Probability Student's True Level Of Functioning Is

Greater Than π_o Given A Uniform Prior Distribution

Minimum Advancement Score	No. of Test Items	Posterior Distribution	Criterion Level--π_o									
			50	55	60	65	70	75	80	85	90	95
6	8	$\beta(7, 3)$	91	85	77	66	54	40	26	14	5	1
7	8	$\beta(8, 2)$	98	96	93	88	80	70	56	40	23	7
8	8	$\beta(9, 1)$	100	100	99	98	96	92	87	77	61	37
7	9	$\beta(8, 3)$	95	90	83	74	62	47	32	18	7	1
8	9	$\beta(9, 2)$	99	98	95	91	85	76	62	46	26	9
9	9	$\beta(10, 1)$	100	100	99	99	97	94	89	80	65	40
7	10	$\beta(8, 4)$	89	81	70	57	43	29	16	7	2	-
8	10	$\beta(9, 3)$	97	93	88	80	69	54	38	22	9	2
9	10	$\beta(10, 2)$	99	99	97	94	89	80	68	51	30	10
8	11	$\beta(9, 4)$	93	87	77	65	51	35	21	9	3	-
9	11	$\beta(10, 3)$	98	96	92	85	75	61	44	26	11	2
10	11	$\beta(11, 2)$	100	99	98	96	92	84	73	56	34	12
9	12	$\beta(10, 4)$	95	91	83	72	58	42	25	12	3	-
10	12	$\beta(11, 3)$	99	97	94	89	80	67	50	31	13	2
11	12	$\beta(12, 2)$	100	100	99	97	94	87	77	60	38	14

that it is unreasonable to require perfect performance as a standard for advancement, and therefore, we need to improve upon this analysis. One way is to use a longer test, but we can, at least, hope to find a procedure in which a twelve-item test will be adequate.

The results in Table 3, although they provide relevant information for mastery decisions about students based on test scores, do not take full advantage of the power which is available through the use of prior knowledge. In particular, it will seldom be the case that our knowledge of a student's true level is adequately described by a uniform distribution. For example, our prior probability that a student is functioning above a criterion level of .8 might be approximately .75. This would be the case if historical data suggested that about 75% of the students who completed a unit of Individually Prescribed Instruction proved to be at or above mastery level. Moreover, we might judge the strength of our knowledge to be roughly equivalent to that based on a score from a 12-item test. (A method for making this assesment will be referenced shortly.)

When working with a binomial model, it is convenient and generally very satisfactory to select a member of the Beta class of distributions to characterize prior beliefs (Novick and Jackson, 1974). If this is done, the posterior distribution is easily obtained, and in every instance will again be a member of the Beta family. In fact, if the prior distribution is $\beta(a, b)$ and x success in n trials are observed, then the posterior distribution is $\beta(x + a, n - x + b)$. This can be seen in Table 3, where it is noted that the uniform distribution is $\beta(1, 1)$. If we restrict ourselves to prior distributions in the Beta family, the beliefs specified in the previous paragraph are characterized by $\beta(10.254, 1.746)$. Given this prior

distribution and the indicated test results, the posterior distributions and posterior probabilities of exceeding various criteria are provided in Table 4. The precise stipulation of prior distributions must always be done carefully, but extensive aids (Novick and Jackson, 1974, Novick, Lewis, and Jackson, 1973) are available, and indeed an elaborate system of Computer Assisted Data Analysis (CADA) is available (Novick, 1973) to help an instructional decision maker specify his prior distribution. A yet more sophisticated way of getting prior and posterior distributions for each person is derived by Lewis, Wang, and Novick (1973) and the required tables are given by Wang (1973). For the present, we shall suppose that this work has been done carefully and that the prior distribution used in the construction of Table 4 is appropriate.

Tables 3 and 4 demonstrate clearly the impact of prior knowledge on our interpretation of test results. In Table 3, for example, the posterior probability that a student with a score of six out of eight items correct has a true level greater than .80 is only .26, whereas in Table 4 this probability has increased to .60. This result should not be surprising, in view of the fact that we have now set this probability to be .75, apriori as compared to .20 in Table 3. If we felt the chances to be very good that the student had mastered an objective (to a level above .8) before we saw the test results, then a score of six out of eight will not substantially change our beliefs; it will lower the probability, but aposteriori may still leave the odds in favor of mastery. In many applications, a prior probability of mastery may be no more than .60, but the results will still differ sharply from those obtained, assuming vague prior information. Note that if we were to adopt the rule that we will advance a student if the aposteriori probability of mastery is at least

Table 4

Probability Student's True Level of Functioning Is

Greater Than π_o Given A $\beta(10.254, 1.746)$ Prior Distribution

Minimum Advancement Score	No. of Test Items	Posterior Distribution	Criterion Level--π_o									
			50	55	60	65	70	75	80	85	90	95
6	8	$\beta(16.254, 3.746)$	100	100	98	96	90	78	60	37	15	2
7	8	$\beta(17.254, 2.746)$	100	100	100	99	97	92	81	62	36	10
8	8	$\beta(18.254, 1.746)$	100	100	100	100	99	98	94	85	66	32
7	9	$\beta(17.254, 3.746)$	100	100	99	97	92	82	65	41	17	2
8	9	$\beta(18.254, 2.746)$	100	100	100	99	98	93	84	66	39	11
9	9	$\beta(19.254, 1.746)$	100	100	100	100	100	98	95	87	69	34
7	10	$\beta(17.254, 4.746)$	100	99	97	93	84	68	47	24	7	1
8	10	$\beta(18.254, 3.746)$	100	100	99	98	93	84	68	45	19	3
9	10	$\beta(19.254, 2.746)$	100	100	100	99	98	95	86	69	42	12
6	11	$\beta(18.254, 4.746)$	100	99	98	94	87	72	51	27	8	1
9	11	$\beta(19.254, 3.746)$	100	100	100	98	95	87	72	48	22	3
10	11	$\beta(20.254, 2.746)$	100	100	100	100	99	96	88	72	45	13
9	12	$\beta(19.254, 4.746)$	100	100	99	96	89	76	55	30	10	1
10	12	$\beta(20.254, 3.746)$	100	100	100	99	96	89	75	52	24	4
11	12	$\beta(21.254, 2.746)$	100	100	100	100	99	96	90	75	48	14

Note: The mean and mode, respectively of $\beta(10.254, 1.746)$ are .855 and .925 and for this distribution Prob($\pi > \pi_o$) for π_o = .70, .75, .80, .85 are .92, .86, .75, and .59, respectively. A close look at these distributional characteristics will help a decision maker determine if this prior distribution is a realistic characterization of his beliefs.

.50, then in this example, we will advance him if the prior distribution were that of Table 4, but not if it were that of Table 3.

When the decision maker specifies an informative prior distribution, he is saying, in effect, that he wants a decision which will have a high probability of being correct in that portion of the decision space in which he thinks the student's ability truly lies. For example, referring to Table 2, a decision maker with a high prior probability that the student had a true level of functioning below .75 would, by virtue of his analysis, require a minimum passing score of seven correct out of eight items. This would assure him a low probability of misclassification for all values below .75. Another decision maker with high prior probability that the student was above criterion level would likely require only six out of eight correct, and thus have low probability of an incorrect decision for values of .75 or above.

Once we have decided to work with the posterior probability that a student's level of functioning exceeds some criterion, given his test score, and have made use of our prior knowledge in obtaining this probability, another issue remains to be settled before we can turn to the question of test length. Simply stated, we need to know how sure we should be that a student has mastered an objective at the chosen level before we make the decision to allow him to advance to the next objective. For instance, is a posterior probability of at least .5, as was used in the last example, a reasonable choice in all cases? Almost certainly this last question should be answered in the negative. The point at issue here comes down to an understanding of the relative disutilities or losses associated with the false positive and false negative errors.

If it were no more serious to advance a student whose level was below the criterion than to retain a student who was above, we would be behaving optimally if we were to advance students with posterior probabilities above .5 and retain the others. In many situations the prior probability will be this high, and hence an advancement decision could then be made on an apriori basis. On the other hand, we might consider the loss to be twice as great for a false advancement than for a false retention. In this case, we should only advance those students whose posterior probability for being above the criterion exceeds 2/3. The general result is that we shall achieve the smallest expected loss if we match the posterior odds to the loss ratio. Thus, if the loss ratio is 2 to 1 (false advance to false retain), a probability of 2/(2 + 1) gives matching odds of 2/3 to 1/3 above criterion to below criterion).

Table 5

Losses Associates With Incorrect Decisions

		True Level	
		$\pi \geq \pi_o$	$\pi < \pi_o$
Decision	Advance	0	a
	Retain	b	0

To express the result symbolically, consider the notation of Table 5. Here a is the loss associated with advancing a student whose true level is below π_o, and b is the loss for retaining a student whose true level exceeds π_o. The decision rule which minimizes expected loss in this situation is

to advance a student if his test score is such that

$$b \ \mathrm{Prob}(\pi \geq \pi_o | x, n) \geq a \ \mathrm{Prob}(\pi < \pi_o | x, n),$$

and to retain him otherwise. This comparison is equivalent to comparing the loss ratio a/b to the probability ratio $\mathrm{Prob}(\pi \geq \pi_o | x, n)/\mathrm{Prob}(\pi < \pi_o | x, n)$.

If a = b in our analysis, the decision procedure reduces to comparing the median of the posterior distribution with the specified criterion level. If the median is at least at this level, the student is advanced, otherwise he is retained. In this situation, the decision procedure is very similar to that used by Millman (1972). Though the procedure used by Millman is not Bayesian, it is equivalent to comparing with the mode (rather than the median) of the posterior distribution based on a uniform prior. Thus, in effect, the sampling theory approach gives equal weight to all equal intervals throughout the range of π; that is, effectively, to take π to be uniformly distributed apriori. This is seldom a reasonable prior specification. We might also remark that the formulation in Table 5 can be generalized to provide for differential utilities for correctly identifying true positives and true negatives as well as differential disutilities (or losses) for false positives and false negatives as is done in Table 5. To do this negative quantities (negative disutilities = utilities) would need to replace the zeros in Table 5, and a slightly more complicated analysis would not be used.

It may be worthwhile to summarize the situation at this point. An instructor wishing to use test results in the context of Individually Prescribed Instruction should be ready to supply three kinds of information. First, a criterion level--the minimum degree of mastery required--must be set. In Individually Prescribed Instruction this seems to run from about

.70 to about .85. Second, prior knowledge of the student's true level of functioning must be translated into probability terms, namely a prior probability distribution for π. Typically, a carefully monitored program will be such as to suggest a prior probability distribution that assigns a probability of just more than .50 to the region above the criterion level. If this is not the case, the general efficacy of the program should be re-evaluated. A program that results in a much higher probability may be wastefully long and one that results in a lower probability may require strengthening. Finally, the relative losses associated with the two types of incorrect decisions must be assessed. A ratio of more than 1/1 is the rule (we are told) with ratios of 1.5/1 and 2/1 being common, and ratios as high as 3/1 not being rare.

It should be clear that all three of the above determinations will have an influence on the minimum necessary test length. As the criterion level approaches unity, the test must be longer in order to provide adequate information about a student's level of functioning in the neighborhood of the criterion. If prior probabilities of mastery are sufficiently high, very short tests become possible, but this is not and should not be the typical case. Finally, higher loss ratios require longer tests to allow the possibility of high posterior probability of mastery. We shall also see that greater test lengths are sometimes required because of the obvious restriction to integer valued sample sizes.

A Design For Test-Length Specification

The characteristics of the group of students being tested must now be considered as they relate to test-length specification. Each member

Table 6

Selected Prior Distributions For IPI Advancement Decisions

No.	Prior Distribution	Effective Prior Sample Size	Mean	Prob($\pi_\ell \le \pi \le \pi_u$)* .00-.70	.70-.75	.75-.80	.80-.85	.85-.90	.90-1.00
1	$\beta(5.6, 2.4)$	8	.70	.46	.12	.12	.12	.10	.08
2	$\beta(6, 2)$	8	.75	.33	.12	.13	.14	.13	.15
3	$\beta(6.4, 1.6)$	8	.80	.21	.10	.12	.15	.16	.26
4	$\beta(6.8, 1.2)$	8	.85	.12	.07	.09	.13	.17	.42
5	$\beta(7.2, .8)$	8	.90	.05	.04	.06	.09	.14	.62
6	$\beta(7, 3)$	10	.70	.46	.14	.14	.12	.09	.05
7	$\beta(7.5, 2.5)$	10	.75	.32	.13	.15	.15	.13	.12
8	$\beta(8, 2)$	10	.80	.20	.10	.14	.16	.17	.23
9	$\beta(8.5, 1.5)$	10	.85	.10	.07	.10	.14	.19	.40
10	$\beta(9, 1)$	10	.90	.04	.03	.06	.10	.16	.61
11	$\beta(8.4, 3.6)$	12	.70	.47	.15	.15	.12	.08	.03
12	$\beta(9, 3)$	12	.75	.32	.14	.16	.16	.13	.09
13	$\beta(9.6, 2.4)$	12	.80	.18	.11	.15	.18	.18	.20
14	$\beta(10.2, 1.8)$	12	.85	.09	.07	.11	.16	.20	.37
15	$\beta(10.8, 1.2)$	12	.90	.03	.03	.06	.11	.17	.60
16	$\beta(10.5, 4.5)$	15	.70	.47	.17	.16	.12	.06	.02
17	$\beta(11.25, 3.75)$	15	.75	.30	.16	.18	.17	.13	.06
18	$\beta(12, 3)$	15	.80	.16	.12	.17	.20	.19	.16
19	$\beta(12.75, 2.25)$	15	.85	.07	.07	.12	.18	.23	.33
20	$\beta(13.5, 1.5)$	15	.90	.02	.03	.06	.11	.19	.59

*Note: All entries have been rounded to two decimal places and smoothed so that the row totals add to 1.00.

of the group of students tested has been exposed to the same instruction program under identical local conditions. If a particular student is not considered atypical for this group, then our prior beliefs about his true level of functioning should closely reflect the true distribution of levels of functioning found in that group. Indeed, elaborate formal procedures for, effectively, bootstrapping a prior distribution using, for each examinee, the scores on the remaining $m - 1$ examinees are described by Novick, Lewis, and Jackson (1973). Thus, a group characteristics, through their effect on our prior distributions, do affect test-length specification. If the average test score of the group is high (i.e., above the criterion level) and there is little variation among individuals, shorter tests become feasible.

Since, in practice, prior distributions will be based upon on-site experience, there will, of course, be different prior distributions for different sites. What we shall attempt to do here is to show what sample sizes will be required for a broad range of prior distributions and loss ratios. What we need to do now, therefore, is to consider certain combinations of prior distributions, criterion levels and loss ratios, and see what sample size will be adequate in each case.

For our analyses, we shall consider 20 different prior distributions for the level of functioning π, four specified criterion levels, and four loss ratios. For each criterion level, we shall consider all four loss ratios and four of the prior distributions. The four loss ratios we shall use are 1.5, 2.0, 2.5, and 3.0. The respective probabilities $P = \text{Prob}(\pi \geq \pi_o)$ required for advancement [given by setting $P/(1 - P)$ equal to the loss ratios, a/b] are .60, .67, .71, and .75. Thus, with a

loss ratio of 3.0, the posterior probability that the student's level of functioning is greater than the specified criterion level must be at least .75, if he is to be advanced.

The twenty prior probability distributions we shall be considering are given in Table 6 where they have been grouped in blocks of five, with each block having a distribution with the respective mean values .70, .75, .80, .85, and .90. The blocks differ with respect to the concentration of the prior distributions. Within block, the distributions differ with respect to their mean values. Note that in the first block the arguments of each Beta distribution sum to 8, e.g., 5.6 + 2.4 = 8. This indicates that the amount of prior information contained in each of these distributions is equivalent to what would be gained from a test containing eight items. If given one of these prior distributions and some criterion level and loss ratio, we specify an eight-item test, our posterior distribution will contain information equivalent to that contained in 16 observations. This contrasts with the classical procedure which uses no prior information. It is this increment in information that is equivalent to prior observations which permits a reduction in test length when a Bayesian procedure is used.

The first problem in doing an analysis is that of selecting a reasonable prior distribution. For the present application, we would first need to ask ourselves what we would expect to find as the mean level of functioning in our posttest group. With a specified criterion level of .70, we might hope for a mean level of functioning of .70. Thus, we would have people in training until such time as we would "expect" them to be qualified. Since loss ratios are typically greater than one, some overtraining may be thought to be useful, but as we shall see, excessive overtraining may be wasteful.

Suppose, for concreteness, that we believe the mean population level of functioning to be .70. Distributions 1, 6, 11, and 16 satisfy this condition, and, hence, we may choose from among these. We note that these distributions are in an increasing order of tightness, as may most conveniently be seen in the probability assignment given in the last column, to the interval (.90, 1.00). These probabilities are respectively .08, .05, .03, and .02. We need to ask ourselves which of these values seems most reasonable, and this then will give us some preference among these prior distributions. We might consider the relative weight of prior information assumed by each prior distribution (8, 10, 12, and 15 equivalent prior observations, respectively), and this should help to narrow our focus to one or two adjacent prior distributions for this, or any other application. Since the authors of this paper cannot know what an appropriate prior distribution will be in applications they have not seen, it will be most helpful, we think, to work out sample size allocations for several prior distributions and leave the final selection to be made "in the field". We believe that the prior distributions, loss ratios, and specified criterion levels used here are typical of those found in practice, and, therefore, that the specific results we shall obtain will be useful. However, if other combinations present themselves, we believe that the general methodology that we are demonstrating should be adequate to the problem. Actually we shall find that most of our specifications are very robust with respect to the choice of prior distribution within the range we have considered.

Some Specific Test Length Recommendations

In Table 7, we give recommended sample sizes and minimum advancement scores for π_o = .70, (a/b) = 1.5, 2.0, 2.5, 3.0 and prior distributions 1, 6, 11, and 16. The values that we have settled on for the body of

Table 7

Recommended Sample Sizes and Advancement Scores

$\pi_0 = .70$

Prior Distribution	$\mathcal{E}(\pi)$	Loss Ratio 1.5 (.60)	2.0 (.67)	2.5 (.71)	3.0 (.75)
$\beta(5.6, 2.4)$[1]	(.70)	6/8(.62)	10/13(.70)	11/14(.74)	12/15(.78)
$\beta(7, 3)$	(.70)	6/8(.61)	10/13(.69)	11/14(.73)	12/15(.77)
$\beta(8.4, 3.6)$	(.70)	6/8(.61)	10/13(.68)	11/14(.72)	12/15(.76)
$\beta(10.5, 4.5)$	(.70)	9/12(.62)[2]	10/13(.67)	11/14(.71)	12/15(.75)
		General Recommendations			
		6/8(75%)	10/13(77%)	11/14(79%)	12/15(80%)

[1] Apriori, Prob($\pi \geq .70$) for each of the four prior distributions is .54, .54, .53, and .53.

[2] For 6/8, Prob($\pi \geq .70$) = .598.

this table are not, in every instance, optimum in any statistical sense, though we are confident that the risks associated with these decision rules are in every case insignificantly different from the risks of the optimum procedures. In selecting values for this table we have sought sample sizes and minimum advancement scores that would be very efficient over a wide range of prior distributions. That we have been successful in this endeavor is confirmed by our ability to give general recommendations that hold throughout the range of prior distributions studied. Actually in only one instance have we cheated (see footnote 2, Table 7), but again the increase in expected loss will be trivial. We would also note that the required percentage correct and the number of required observations increases as the loss ratio increases, which "makes sense" on intuitive grounds.

A rough indication of the near optimality of any of the individual specifications can be gained from the closeness of the aposteriori probability (indicated in parentheses following the specification) with the value required by the particular loss ratio (given in parentheses at the top of the column). Thus, with the prior distribution $\beta(7, 3)$, the decision rule "six out of eight", abbreviated 6/8, leads to the aposteriori distribution $\beta(13, 5)$ and to $\text{Prob}(\pi > .70) = .61$ which is just .01 greater than the required level .60 for the loss ratio 1.5 (1.5 to 1). In this instance, the specified decision rule may be very good. On the other hand, consider the prior distribution $\beta(5.6, 2.4)$. Here the rule 11/14 leads to a value .74 when only .71 is required for a 2.5 to 1 loss ratio.

Actually, the specification 8/10 is somewhat better giving a posterior probability of .729. Also for the prior distribution $\beta(7, 3)$, the posterior probability with 8/10 is .718. With the loss ratio 2.0/1 and with the prior $\beta(5.6, 2.4)$, the rule 7/9 leads to the posterior probability .68 as compared to desired value of .67. In every case where we have specified an "almost best" decision rule, the result has been an increase in the specified sample size and the purpose has been to obtain uniformity of specification over a reasonably wide range of amounts of prior information. Considering our general ignorance concerning what might be an appropriate prior distribution in specific applications, the specifications we have given should be the more generally useful.

Another indication of how good a particular specification is can be inferred from the closeness of the percentage correct required by the advancement rule to the specified criterion level. Clearly, if the percentage required by the advancement rule is very much larger than the specified criterion level, a large percentage of qualified students will be retained and this is undesirable, particularly for small loss ratios. For large loss ratios, this is less important and hence higher advancement ratios can, and will need to be tolerated. This feature is exhibited in Table 7, where the advancement ratios increase with increasing loss ratios. One can, of course, keep the advancement ratio down very close to the specified criterion level even for higher loss ratios, but only by having much larger sample sizes. For example with the prior distribution $\beta(5.6, 2.4)$ the specified criterion level $\pi_0 = .70$ and the loss ratio 2.0, the advancement ratio 72/100 is satisfactory since $\text{Prob}(\pi > .70 | 72/100) = .675$, but the indicated sample size is unacceptable.

Table 8

Recommended Sample Sizes and Advancement Scores

$\pi_o = .75$

Prior Distribution	$\mathcal{E}(\pi)$	Loss Ratio 1.5 (.60)	2.0 (.67)	2.5 (.71)	3.0 (.75)
$\beta(6, 2)$[1]	(.75)	8/10(.65)	16/20(.70)	17/21(.74)	18/22(.77)
$\beta(7.5, 2.5)$	(.75)	8/10(.64)	16/20(.69)	17/21(.73)	18/22(.76)
$\beta(9, 3)$	(.75)	8/10(.63)	16/20(.69)	17/21(.72)	18/22(.75)
$\beta(11.25, 3.75)$	(.75)	8/10(.62)	16/20(.68)	7/21(.71)	19/23(.77)[2]
		General Recommendations			
		8/10(80%)	16/20(80%)	17/21(81%)	18/22(82%)

[1]Apriori, Prob($\pi \geq .75$) = .56, .55, .55, and .54, respectively, for the four prior distributions used in Table 8.

[2]For 18/22, Prob($\pi \geq .75$) = .744.

Table 9

Recommended Sample Sizes and Advancement Scores

$\pi_o = .80$

Prior Distribution	$\mathcal{E}(\pi)$	Loss Ratio 1.5 (.60)	2.0 (.67)	2.5 (.71)	3.0 (.75)
$\beta(6.4, 1.6)$[1]	(.80)	6/7(.66)	7/8(.70)	17/20(.72)	19/22(.78)
$\beta(8, 2)$	(.80)	6/7(.65)	7/8(.69)	17/20(.72)	19/22(.77)
$\beta(9.6, 2.4)$	(.80)	6/7(.64)	7/8(.68)	17/20(.71)	19/22(.76)
$\beta(12, 3)$	(.80)	6/7(.63)	7/8(.67)	18/21(.73)[2]	19/22(.75)
		General Recommendations			
		6/7(86%)	7/8(88%)	17/20(85%)	19/22(86%)

[1]Apriori, Prob($\pi \geq .80$) = .57; for 8/10, Prob($\pi \geq .80$) = .55; for 16/20, Prob($\pi \geq .80$) = .54; for 8.5/10, Prob($\pi \geq .80$) = .67; for 8.3/10, Prob($\pi \geq .80$) = .62; for 9/10, Prob($\pi \geq .80$) = .78.

[2]For 17/20, Prob($\pi \geq .80$) = .70.

Note that for each of the prior probabilities used in Table 7, $\text{Prob}(\pi \geq .70) > .50$. Thus, on an apriori basis, advancement would be indicated with a loss ratio 1.0. This will generally be true for the prior distributions we shall be adopting for our analyses. The point is that loss ratios of 1.0 are not (we are told) typical of IPI applications, and if test lengths are to be kept reasonable it will be necessary to use training programs that give mean output at or above the criterion level.

There has been a definite tendency in IPI to require relatively high advancement ratios; typically, the value .85 is used. One might well speculate whether this is a function of a high loss ratio combined with a desire for a short test length, or whether it really reflects a perceived need for a high criterion level. (For example an advancement ratio of 6/7 with the prior distribution $\beta(5.6, 2.4)$ would yield with $x = 6$ a posterior $\text{Prob}(\pi > .70) = .77$ which would be just right with a loss ratio of 3.0.) The authors of this paper do not know the answer to this question, but hope that those within IPI will want to consider it carefully. Only through such serious consideration can the test length problem be "solved".

Some recommended test lengths for $\pi_o = .75$ and four prior distributions with $\mathcal{E}_\rho(\pi) = .75$ are given in Table 8. Again we have been able to specify one generally satisfactory advancement ratio for each of the four loss ratios. We note that the required test lengths for $\pi_o = .75$ are rather larger than for $\pi_o = .70$. In Table 8, we find very short required test lengths for a 1.5 loss ratio and rather long ones for loss ratios of 2.0, 2.5, and 3.0.

In Table 9, we provide recommendations for $\pi_o = .80$ when $\mathcal{E}_\rho(\pi) = .80$. The results here parallel those of Table 8, except that the advancement ratios are very high as compared to the criterion levels. This is

relatively unsatisfactory. In Footnote 1 to Table 9, we indicate the formal results for the prior distribution $\beta(6.4, 1.6)$ and the sample result "8.5" correct and "1.5" incorrect and also for "8.3" correct and "1.7" incorrect. These provide very nice results for loss ratios of 2.0 and 1.5, respectively. Unfortunately, these are unobtainable sample results. This demonstrates that in part, large required test lengths may sometimes be due to the discreteness, and hence, discontinuity of our possible experimental outcomes. This also suggests that the precise specification of the advancement rules may be highly sensitive to the mean value of the prior distribution even if it is proving to be relatively insensitive to the total amount of information contained in the prior distribution, which is indicated by the sum of the two parameters of the Beta distribution.

For example, given the prior distribution $\beta(6.4, 1.6)$ and the impossible sample result $x = 8.3$, $n = 10$, we have the posterior distribution $\beta(14.7, 3.3)$ which, as we indicated previously, gives $\text{Prob}(\pi > .80) = .62$ which suggests that the advancement ratio 8.3/10 might be very favorable with a loss ratio of 1.5. But suppose we had just a slightly different prior distribution, namely, $\beta(6.7, 1.3)$ with $\mathcal{E}(\pi) = .84$, then the sample result $x = 8$, $n = 10$ would yield the posterior distribution $\beta(14.7, 3.3)$ and thus, for the reasons given above, indicate that the advancement ratio 8/10 might be attractive. This advancement ratio is clearly more attractive than the ratio 6/7, despite the fact that it requires three additional items, because this ratio 8/10 = 80% is closer to the criterion level than is the advancement ratio 6/7 = 86%.

Because of this relatively high dependence of the results on the expected value of the prior distribution, it seems important to attempt some study of the variation of our results as a function of changes in

Table 10

Recommended Sample Sizes and Advancement Scores

$\pi_o = .80$

Prior Distribution	$\mathcal{E}(\pi)$	Loss Ratio 1.5 (.60)	2.0 (.67)	2.5 (.71)	3.0 (.75)
$\beta(6.8, 1.2)$[5]	(.85)	8/10(.64)	9/11(.69)	10/12(.72)[1]	11/13(.76)
$\beta(8.5, 1.5)$	(.85)	8/10(.66)	9/11(.70)	10/12(.73)[2]	11/13(.76)
$\beta(10.2, 1.8)$	(.85)	8/10(.67)	9/11(.71)	9/11(.71)[3]	11/13(.77)
$\beta(12.75, 2.25)$	(.85)	8/10(.69)	9/11(.72)	9/11(.72)[4]	11/13(.78)
			General Recommendations		
		8/10(80%)	9/11(82%)	10/12(83%)	11/13(85%)

[1]For 5/6, Prob($\pi \geq .80$) = .72.

[2]For 5/6, Prob($\pi \geq .80$) = .73.

[3]For 10/12. Prob($\pi \geq .80$) = .74.

[4]For 10/12, Prob($\pi \geq .80$) = .75.

[5]For the four prior distributions, the apriori probabilities of $\pi \geq .80$ are .72, .73, .74, and .75. With these prior distributions and with $7/\overline{10}$, the posterior probabilities of $\pi \geq .80$ are .41, .43, .46, and .48.

our prior distribution. For this reason, we have in Table 10 redone our sample size recommendations under the assumption that the mean of our prior distribution is .85 instead of .80.

Surely the practitioner will find the sample size recommendations of Table 10 to be attractive. Apparently with these prior disbributions, test lengths need be no greater than 13 for any of the listed loss-ratios. With the prior distributions having $\mathcal{E}(\pi) = .80$, a sample size of 22 is required when the loss ratio is 3.0.

What is happening is that we are beginning with fairly strong beliefs that $\pi \geq \pi_0$ so that not much data, in confirmation, is required even for high loss ratios. In fact, even on an apriori basis, an advancement decision would be made for all loss ratios up to and including 2.5. Indeed, we see that the function of the sample data here is to provide the possibility of obtaining some information that might change the decision to retention. For example, an observed performance ratio of 10/13 with the prior distribution $\beta(6.8, 1.2)$ would give aposteriori $\text{Prob}(\pi \geq .80) = .72$, and hence, the student would be retained if the loss ratio were 3.0 (see also Footnote 5, Table 10).

We believe that the comparison of the specifications in Tables 9 and 10 have important implications for IPI management. When loss ratios are high, it may well be highly advantageous to strengthen the training program to the extent that the mean output is well above the specified criterion level. This will make it possible to use short tests or. alternatively will generally reduce the risk of incorrect classification. This will, of course, be more expensive, and this investment must be balanced out against the reduction in the cost of testing and the reduction in the expected loss due to incorrect decision. The final Table, Table 11, looks

Table 11

Recommended Sample Sizes and Advancement Scores

$\pi_o = .85$

Prior Distributions	$\mathcal{E}(\pi)$	Loss Ratio 1.5 (.60)	2.0 (.67)	2.5 (.70)	3.0 (.75)
β(6.8, 1.2)[1]	(.85)	7/8(.62)	9/10(.70)	17/19(.73)	18/20(.76)[3]
β(8.5, 1.5)	(.85)	7/8(.62)	9/10(.69)	17/19(.72)	19/21(.77)
β(10.2, 1.8)	(.85)	7/8(.61)	9/10(.68)	17/19(.72)	19/21(.76)
β(12.75, 2.25)	(.85)	7/8(.60)	9/10(.67)	17/19(.71)[2]	19/21(.75)
		General Recommendations			
		7/8(87.5%)	9/10(90%)	17/19(89%)	19/21(90%)

[1]The apriori probabilities for $\pi \geq .85$ are .59, .58, .58, and .57.

[2]For 10/11, Prob(π > .85 = .695).

[3]For 19/21, Prob(π > .85 = .78).

very much like Table 9 as far as test lengths are concerned. Here again some robust length assignments are obtained, though again, the lengths for the high loss ratios border on being discomforting. This can be corrected by training to an average level of functioning of .90. With the prior distribution $\beta(7.2, 8)$, we find that $\text{Prob}(\pi \geq .85) = .76$, apriori. Observing 6/7 yields $\text{Prob}(\pi \geq .85) = .70$, while 5/7 yields a value of .41. Observing 8/9 yields .77, while 7/9 yields .493. Thus, clearly, very short test lengths are again possible if the students are trained to a sufficiently high average standard.

Some Summary Remarks

The test length recommendations given in this paper are meant to be taken seriously and hopefully they will soon be adopted on a provisional and experimental basis, so that more experience can be gained while some of the theoretical and substantive issues raised in this paper are debated. The questions of level of functioning required to define mastery and the relative losses incurred in making false positives and false negative decisions require serious discussion and concensus. We also need to get some clear picture of what kinds of distributions of outcomes are to be expected as this determines the amount of prior information available in making individual assessments. This third issue is, as we have indicated, intimately related to the expected level of functioning that is sought in the group being trained. Hopeful and possible outcomes of such discussions could be a consensus that:

1. In most situations a level of functioning of something less than .85 is satisfactory. A value as low as .75 would be highly desirable. This could be accomplished by redefining the task domain slightly to eliminate very easy items.

2. Training should be carefully monitored so that expected group performance will be just slightly higher than the specified criterion level. This will keep training time and testing time relatively low.

3. The program should be structured so that very high loss ratios are not appropriate. That is to say, individual modules should not be overly dependent on preceding ones.

One problem that does not arise with Bayesian methods is any complication if sequential methods are used. Items can simply be administered until it is clear that a student will definitely, or cannot possibly, attain the minimum advancement score. Thus with a minimum advancement score of 8/10, testing can cease as soon as light successes or three failures are observed.

Two issues have been treated in a rather gross way in this paper and on these important issues further research needs to be done. First it must be recognized that while the threshold loss function we have adopted here is a better approximation to reality than, for example, Livingston's criterion centered squared-error loss (see Hambleton and Novick, 1973), it is only a gross approximation to be used while better and more complicated approximations are being investigated. Three that immediately come to mind are:

1. A threshold loss function with an indifference region in which there is zero loss for false positive or false negative errors.

2. A <u>negative</u> squared-exponential loss used with the root arcsine transformation parameter

$$\gamma = \sin^{-1}\sqrt{\pi} \quad .$$

3. A cumulative Beta distribution loss function.

We expect that these loss functions will give somewhat different and surely better length specifications than those obtained here, but the overall decrease in expected loss may or may not be great. We should also remark that these recommendations are specifically made for first time through decisions. We have yet to consider the problem of decisions for students repeating a unit.

Finally, we would remark that one of the important issues that we identified at the outset of this paper has been handled in a most casual and informal manner. To do other than this would have enormously complicated the analysis and delayed substantially the appearance of our recommendations. We refer explicitly to the premium on testing time within the instructional process and implicitly to an implied trade-off between training and testing time. A completely general analysis would consider an available time T and an allocation of T into instruction and testing times $i + t = T$, so as to maximize a payoff function which would have a (possibly differential) positive payoff for each module successfully completed, and a (differential) negative payoff for an incorrect decision of either type. We are reluctant to undertake such a sophisticated analysis until such time as the operating conditions of IPI are more clearly defined.

For the present paper we have implicitly adopted some guidelines which effectively say that it is very desirable to have test lengths of 12 or less, tolerable but undesirable to have test lengths as high as 20 and discomforting to have tests that are longer than this. We have also taken the position that a decision should not be made on the basis of prior and

collateral information alone but that mastery must be confirmed by a test that permits demonstration of nonmastery. As in all of the judgmental decisions made in this paper we have been guided by counsel from experienced IPI personnel, particularly Richard Ferguson and Anthony Nitko to whom we are much indebted. The value of this paper will largely be determined by the quality of the discussion engendered by it among such people.

REFERENCES

Hambleton, R., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973.

Lewis, C., Wang, M., & Novick, M. R. Marginal distributions for the estimation of proportions in m groups. ACT Technical Bulletin No. 13. Iowa City, Iowa: The American College Testing Program, 1973.

Millman, J. "Determining Test Length, Passing Scores and Test Lengths for Objectives-Based Tests." Los Angeles: Instructional Objectives Exchange, 1972.

Novick, M. R. High school attainment: An example of a computer-assisted Bayesian approach to data analysis. International Statistical Review, 1973, 41, 264-271.

Novick, M. R., & Jackson, P. H. Statistical Methods for Educational and Psychological Research. New York: McGraw-Hill, 1974.

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-46.

Wang, M. Tables of constants for the posterior marginal estimates of proportions in m groups. ACT Technical Bulletin No. 14. Iowa City, Iowa: The American College Testing Program, 1973.

The article, "Toward An Integration of Theory and Method for Criterion-Referenced Tests" has been removed from this document due to copyright restrictions. The entire article appears in Journal of Educational Measurement; Vol. 10, No. 3, Fall 1973.

ACT RESEARCH REPORT

No. 60

60

September 1973

IMPLEMENTATION OF A BAYESIAN SYSTEM FOR DECISION ANALYSIS IN A PROGRAM OF INDIVIDUALLY PRESCRIBED INSTRUCTION

R. L. Ferguson
M. R. Novick

PUBLISHED BY THE RESEARCH AND DEVELOPMENT DIVISION

THE AMERICAN COLLEGE TESTING PROGRAM

P. O. BOX 168, IOWA CITY, IOWA 52240

ABSTRACT

The decision process required for Individually Prescribed Instruction (IPI), an adaptive instructional program developed at the University of Pittsburgh, is described. In IPI, short tests are used to determine the level of proficiency of each student in precisely defined learning objectives. The output of these tests is used to guide instructional planning for individual students.

The nature and effect of errors in proficiency decisions are described and a procedure for reducing the probability of such errors is proposed. The plan calls for a Bayesian procedure which would incorporate prior information on the instructional program, for example the distribution of the percentage of items answered correctly by students Such a procedure would permit inferences about the true level of functioning of each student.

The final section of the paper proposes two methods for implementing these procedures in an ongoing IPI program. one approach calls for the integration of the procedure as a part of a computer-based instructional management system, whereas the second approach describes how the procedure can be made tractable in a typical. non-automated IPI classroom

Prepared by the Research and Development Division
The American College Testing Program

For additional copies write:
Publications Division
The American College Testing Program
P.O. Box 168, Iowa City, Iowa 52240

(Check or money order must accompany request.) Price: $1.00

IMPLEMENTATION OF A BAYESIAN SYSTEM FOR DECISION ANALYSIS IN A PROGRAM OF INDIVIDUALLY PRESCRIBED INSTRUCTION

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U.S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal investigator Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

IMPLEMENTATION OF A BAYESIAN SYSTEM FOR DECISION ANALYSIS IN A PROGRAM OF INDIVIDUALLY PRESCRIBED INSTRUCTION

Richard L. Ferguson
Melvin R. Novick

INTRODUCTION

The feasibility of instructional programs designed to adapt to the individual needs of learners has been adequately demonstrated by educational systems like Individually Prescribed Instruction (Glaser, 1968) and A Program for Learning in Accordance with Needs (Flanagan, 1967). Although these programs accomplish individualization in somewhat different ways, each includes components which can be described by the following sequence of operations:

1. Specification of the learning objectives in terms of observable student behavior
2. Assessment of the student's entering competencies.
3. Assignment or election of educational materials and/or experiences fitted to the student's individual needs.
4. Continuous assessment and monitoring of the student's performance and progress

Since programs like IPI and PLAN call for adaptation of the learning environment to meet individual requirements, they necessarily rely heavily on the systematic assessment of student progress. Glaser (1968) has observed that, in IPI, test data serve as the primary source of information enabling teachers to make differential decisions regarding student instruction. Thus, steps (2) and (4) play a prominent role in the successful implementation of IPI. A review of current decision-making procedures for four selected individualized instructional programs has been given by Hambleton (1973).

The fundamental purpose for testing in individualized instructional programs like IPI and PLAN is to ascertain whether or not the student has attained some prescribed level of proficiency in a specified learning objective. Hambleton and Novick (1973) have observed that, "Questions of precise achievement levels and comparisons among individuals on these levels seem to be largely irrelevant." Because test data are used initially to place a student at the appropriate point within an instructional program or sequence, and thus to identify appropriate learning materials or experiences given his needs, the test models which have emerged to serve this function are very different from those used for standard instructional models. Because these tests relate a student's performance on items drawn from a carefully specified domain to a prespecified criterion or standard, these tests have come to be called *domain* or *criterion-referenced tests*.

It is not the purpose of this paper to contrast the differences between norm-referenced tests and criterion-referenced tests. Suffice it to say that criterion-referenced tests are deliberately con-

71

structed so as to yield measurements which are directly interpretable in terms of specified performance standards (Glaser and Nitko, 1971). The process of constructing such tests involves the specification of a domain of tasks that the student should be able to perform and the selection of samples of these tasks representative of that domain. The student's competency in the skill is judged in terms of his performance in responding to the sample of the tasks which is drawn. Performance on this sample is used to infer that his *level of functioning* in the domain either does or does not meet some prescribed standard.

Because student performance on tests used in IPI and PLAN is used as the basis for making decisions affecting placement and advancement, and because it is crucial that these decisions be accurate, major importance is attached to the precision with which each person's true domain score (level of functioning) can be related to the prescribed proficiency level However, due to time constraints, the tests are often comprised of a very small number of items, usually 10 or less. Thus, the precision of judgment from such tests must be open to question. Because of the important role which testing plays in the instructional decision making within IPI, improvement in the quality of the decision process would be greeted with considerable enthusiasm if it could be accomplished without a corresponding increase in the length of the tests. This paper is addressed to the problem of showing precisely how some new developments in statistical theory make this goal attainable. More specifically, the present paper indicates precisely how these Bayesian methods could be integrated into an ongoing IPI program. In order to lay a proper foundation, one describing the exact nature of the measurement problem in IPI, we propose to confine discussion to one major component of the system, the mathematics program. To this end, a general description of the assessment instruments used in IPI mathematics is contained in the next section.

The mathematical and statistical models which form the basis of the proposed application, and the outline of this application, are based on the work of Novick, Lewis, and Jackson (1973), and the amplifications contained in Lewis, Wang, and Novick (1973), Wang (1973), and Wang and Lewis (1973a, 1973b). A theoretical discussion of these methods is contained in Hambleton and Novick (1973). The Bayesian methods of statistical inference developed in these papers combine direct observation information on each student with certain background information, to permit more accurate decision-making than would be possible without the use of this background information. The use of this background information makes possible the gain in accuracy without additional testing.

THE IPI MATHEMATICS PROGRAM

Ferguson (1970a) provides a detailed description of the IPI Mathematics program. Highlights of that description are provided in subsequent parts of this section In particular, attention is given both to the structure of the curriculum and to the test model which plays such an important role in the management of the program.

The Curriculum

Figure 1 conveys the general organization of the mathematics curriculum. Ten content areas, Numeration/Place Value, Addition/Subtraction, Multiplication, Division, etc., are identified; each occurring at various *levels* of difficulty. The ten areas are listed in a hierarchical order that is followed in instruction The intersection of each level with a specific content area determines a *unit* that consists of a set of behaviorally defined *objectives* or *skills*. Each number in the table indicates the number of skills in the unit. Thus, E level-Systems of Measurement is a unit that consists of a set of five behavioral objectives (skills) which share a similar content but are less difficult than the skills contained in the F level-Systems of Measurement. The absence of a number at any position in the chart indicates that no unit exists for the corresponding content area and level. At the bottom of Figure 1, we have listed the specific behavioral objectives for E level-Systems of Measurement.

The Test Model

As previously indicated, the assessment instruments in IPI perform a dual role in the program, serving both a placement and a diagnostic function. The tests are placement oriented in the sense that they locate a student's position in the curriculum with respect to the skills for which he lacks sufficient proficiency, but for which he has the necessary

	Level						
	A	B	C	D	E	F	G
Numeration/Place Value	15	9	14	5	6	7	6
Addition/Subtraction	17	12	13	10	4	4	6
Multiplication		4	7	9	7	4	3
Division		3	4	7	9	5	6
Fractions	3	3	6	7	11	8	8
Money	1	1	5	5			
Time		6	6	4	4	2	
Systems of Measurement		3	6	6	5	5	6
Geometry		3	2	4	6	4	2
Applications		3	8	9	5	4	6

Behavioral Objectives

E Level-Systems of Measurement

1. Given a ruler, the student measures a line segment with the indicated degree of precision LIMIT: smallest unit of precision 1/8 inch; line segments to 10 inches.

2. Given 20 cut-out regions that are each 1-inch squares and an illustration of a rectangular region, the student uses the 1-inch squares to determine the area of the given rectangular region LIMIT: areas < 20 square inches. Length of sides of rectangles must be multiples of 1 inch.

3. Given the measures of the sides of a rectangular region, the student determines the area of that region LIMIT: integral measures, one unit of measure per problem; units of measure—inches, feet, yards, miles

4. Given the measure of the sides of a rectangular region, the student determines the perimeter and the area of that region. LIMIT. At least one of the measures (length, width) must be integral; both measures must be < 100; one measure may be a common fraction < 1 with denominator < 10; 1 unit of measure per problem; units of measure—inches, feet, yards, miles.

5. Given a weight measurement, the student completes a statement to show an equivalent measurement in a different unit of weight measure. Given a word problem that requires conversion of a given weight measurement expressed in standard units to an equivalent weight expressed in another standard unit, the student solves the problem and writes the answer with the appropriate label LIMIT: units—ounces, pounds, tons.

Fig. 1. Matrix of Units in the IPI Mathematics Curriculum.

prerequisite skills so that he can begin work. The same tests are diagnostic in that they provide information that identifies skills in which the student has not achieved sufficient proficiency and also provide insight as to specific facets of these skills on which instruction is required. A review of the various tests utilized in the mathematics program follows.

Curriculum Placement Tests

Upon entrance to the mathematics program, the placement tests provide a global picture of each student regarding his level of proficiency with respect to the skills in each unit of the curriculum. The data generated by the placement tests are used to develop a profile for each student indicating those units in which he has sufficient proficiency in all of the skills and those in which he has insufficient proficiency For example, the outcome of a placement test might yield a profile indicating sufficient proficiency in all of the skills in level D of the curriculum, and insufficient proficiency in the skills of units at a higher level of difficulty. In this case, the student would begin work in units at level E of the curriculum More typically, a student might demonstrate proficiency at level D-Numeration/Place Value, level F-Addition/Subtraction, level E-Money, level C-Time, and perhaps level D in all other areas Such a student would probably then begin instruction in level C-Time, this being the lowest level in the area hierarchy at which instruction is prescribed.

Because of the global nature of placement tests, they must assess a very large domain of mathematics skills Consequently, practicality demands that the tests include only a small number of items on key objectives in each unit of the curriculum. Thus, important placement decisions are necessarily dependent on tests with a small number of items.

Unit Pretests

Once a placement test has been used to determine a profile for a student, a decision can be made, as indicated in the previous section, regarding the unit on which the student begins his work. At this point, a unit pretest is administered to identify the specific objectives in the unit for which the student has sufficient (insufficient) proficiency. Each pretest consists of several short subtests, one for each objective in the unit.

It is possible for a student to demonstrate sufficient competency on all objectives in the unit. If this were to occur, the student would continue working at the same level, but proceed to the next unit in the area hierarchy where he would be given another unit pretest. Thus, the pretest provides additional information about a student, information which is focused at the level determined by the placement test.

The pretest decision can and sometimes does override a part of the placement decision. This occurs when proficiency is demonstrated by the student in areas and at levels not indicated by the placement test. Thus, the IPI testing paradigm initially involves a two stage semisequential testing program with the placement test largely determining the level at which more intensive testing is to take place.

After the unit pretest has identified the specific skills for which the student requires instruction, student test performance on each of these objectives is examined by the teacher to identify particular types of errors or patterns of errors. In this manner, learning materials and/or experiences consonant with the individual's needs can be prescribed.

The typical pretest includes between six and (preferably) ten items for each objective. Obviously, the size of the domain of items varies with the particular skill. Usually, however, the domain is quite large. Thus, important instructional decisions are often based on student performance on a small number of items that have been representatively sampled from a very large domain. The relative shortness of the tests can certainly be justified from a practical point of view. Longer tests might be considered repressive and would certainly exceed reasonable bounds in terms of the proportion of time given over to them within the total instructional process. Thus, it would appear that the key to more effective and more reliable decisions lies not in increasing the length of the tests beyond, say, eight or ten items, but rather in making better use of the data available within the present system.

Curriculum Embedded Tests

These short "quizzes" measure the student's level of proficiency in a single skill within the curriculum. The written instructional material for each skill in a mathematics unit contains two curriculum embedded tests (CETs). The tests are self-evaluation devices used by the student as a check on his progress as it relates to his work on a given skill. Thus, the student who has completed several learning activities related to the development of his proficiency in a particular skill might take a CET to determine whether he has attained sufficient

proficiency at this point or whether he needs to complete additional steps in the instructional process.

The CET typically consists of from four to six items. Because these short tests serve primarily as self-checks for the student, and because no crucial instructional decision is dependent upon student performance on these tests, they seem to be adequate for the task which they serve.

Unit Posttests

These instruments are equivalent forms of the unit pretests. They are generally administered after the student has concluded learning activities for all skills for which he was identified as being insufficiently proficient on the unit pretest. On the basis of the student's performance on the posttest, he is either advanced to the next unit or required to work with additional instructional materials on those skills for which his test performance did not indicate that he achieved a sufficient level of proficiency. A student generally does not advance to a new unit until he has demonstrated sufficient proficiency for all objectives of the current unit.

As with the pretests, decisions resulting from an analysis of posttest data rely upon tests which generally contain a small number of items. Because incorrect proficiency decisions can be detrimental to the student's progress, a procedure which could add substantially to the accuracy of the decision without increasing the length of the test would be most worthwhile.

THE INSTRUCTIONAL DECISION PROCESS

In this section, the process by which test data are used to make instructional decisions is briefly summarized. In addition, a discussion of the nature and consequences of decision errors resulting from the analysis of test data is presented.

A Summary of the Decision Process

Gross placement tests which sample a broad cross section of the important skills in each unit of the mathematics curriculum are administered upon each student's entry into the IPI program. Score data resulting from these tests are used to determine a profile suggesting the student's level of proficiency in each content area of the curriculum.

At this point, the student completes a pretest for the first unit in the curriculum continuum in which his level of proficiency is insufficient. The profile resulting from the pretest identifies those skills for which learning materials and/or experiences are required if the student is to achieve the specified level of performance. During the instructional process, curriculum embedded tests are available to the student as a means of self-evaluation and an estimate of progress as he works on the skills. After he has completed work on all skills in the unit and is satisfied that he has sufficient competency in all of the unit skills, he is administered a posttest which verifies his progress or identifies those skills for which additional instruction is indicated. Once the unit is successfully completed, the student advances to the next unit on his prescription where he is administered a pretest and the cycle is repeated.

The Nature and Effect of Decision Errors

The placement tests, pretests, and posttests are used primarily to verify that a student either has sufficient proficiency, i.e. mastery, in a given set of skills or that he has an inadequate level of proficiency in those skills. Clearly, it is desirable that the mastery decisions for a student be as accurate as possible. The importance of accuracy of the mastery decision for a student is perhaps best emphasized by a discussion of the consequences of an incorrect decision.

As previously indicated, the IPI tests are constructed by sampling items from the domain of items for the objectives included on the tests. Since any sampling which does not exhaust the population of items for an objective can lead to an incorrect mastery decision and since exhaustive testing is impossible, it is necessary to tolerate the risk of making wrong decisions. In an IPI context, a Type I (α) error occurs when an examinee has sufficient proficiency in a skill but the outcome of the testing suggests that he does not. As a result, he is prescribed work lessons which may serve no significant function. A Type II (β) error occurs whenever the examinee, in fact, lacks proficiency in a skill but on the basis of test results is said to have sufficient proficiency. The consequence of a Type II error is that needed remedial instruction is not provided. A Type II error is perceived to be potentially more serious than a Type I error since the Type II error could easily result in the student having difficulty proceeding through a unit and might

eventually lead to an impasse in instruction; whereas, the Type I error will at worst require that the student pursue a review-like study of skills in which he is already proficient.

Although it is clear that the magnitude of the consequences of an incorrect proficiency decision for a student varies with the direction of the error, it is equally clear that in both cases the error may have detrimental effects for the student. The fact that the tests on which these decisions are based have a small number of items per skill suggests that such errors probably occur quite frequently. Given the constraints imposed by a program which already has a heavy testing component, increasing the length of the tests is not a tractable method for achieving increased accuracy in the mastery decision process. However, it may very well be possible to incorporate additional information into the decision process and thus improve the overall accuracy of the decisions being made. It is this hypothesis to which the remainder of this paper is addressed.

In IPI, as in all individualized instructional programs, decisions are focused around the individual student. If a statistical procedure that uses information other than that contained in the immediate direct observations on the student is contemplated, then a Bayesian procedure incorporating prior information on each student comes to mind. This information would consist of results of the student's performance on previous instructional units. In this way, interindividual variability on prior test performance would be helpful in making current decisions.

The problem with this thinking is that the entire thrust of individualized instruction works toward a reduction of interstudent variability of test results. A student moves ahead to a new unit of instruction only when, it is thought, he is prepared to do so. Indeed, he is encouraged not to take the unit posttest until there is strong evidence that he is prepared to perform well on it. A great deal of posttest score variability is in fact observed, but much of it, though not all, results from unreliability due to the necessarily short length of these tests. Thus, realistically, there is little or no useful differential prior information about the individual student.

On the other hand, there is a great deal of information available about the instructional program. Quite specific information is available concerning the distribution of the percentage of items answered correctly by students (Novick, Lewis, and Jackson, 1973), and it is thus possible to make inferences about the true level of functioning of each student, and the mean and standard deviation of these true values in the population of students. Of course, if the instructional programs were completely efficient and the students were without human frailties, there would be *no* variation in *true* levels of functioning of students on posttests. A student would remain in a unit only until that instant at which his level of functioning attained the prespecified criterion. Nothing approaching this is possible with present instructional technology. However, if we knew this were the true state of affairs, then we would ignore individual test scores and use our information on the group mean and variance to make a positive proficiency decision for all students.

In the real world of Individually Prescribed Instruction there will be some variation in true levels of functioning among students on posttests. The delicate manner in which background information is combined with the direct observational data in the Bayesian decision process, and the increment in decision-making accuracy resulting therefrom is detailed in Novick, Lewis, and Jackson (1973) and Lewis, Wang, and Novick (1973).

Finally, we may note one additional source of background information that can be utilized when, as in IPI, testing involves joint measurement on several skills, simultaneously. In this situation and assuming some relationship among the skills, it is possible to use the collateral information contained in the $t - 1$ of t tests scores for each person to help estimate each t-th test score. Thus, if a person scored highly in $t - 1$ subtests and a little less highly in the t-th, we would suspect that this might be due *in part* to bad luck or carelessness, and we would be inclined to make some adjustment in our estimate of his proficiency on that skill. The Bayesian theory and methods described by Wang and Lewis (1973a, 1973b) provide the rationale and prescription for doing this.

Implementation Procedures

The decision analysis procedures employed by teachers and students in the IPI program must not be overly complex. Thus, the final output of the data analysis procedures used to judge the level of proficiency of a student must be so simple that teachers, aides, and even *students* can read the results, interpret them, and then take whatever action is indicated. It will be permissible to use sophisticated statistical methods, but teachers, aides, and students must not be required to understand much more than is contained in this

paper. In short, although it is not necessary that teachers and students understand the details of the analysis, they must be provided information which facilitates their instructional decision making. In the following section, procedures for dealing with the preceding concerns are discussed.

The collection and analysis of data. During the past several years, considerable investigation has been underway into the feasibility of using a computer as an integral part of the IPI program. A thorough discussion of the most recent developments is available in a progress report (Block, Carlson, Fitzhugh, et al., 1973) recently released by the Learning Research and Development Center at the University of Pittsburgh. Earlier reports include Cooley and Glaser (1969), Ferguson (1970b, 1971), and Ferguson and Hsu (1971). The activities described in these reports emphasize somewhat visionary ideas for how the computer can best be employed in an Individualized program of instruction. Although these studies include the more conventional modes of computer-assisted instruction, they extend far beyond into such areas as computer testing and instructional management.

It is in this latter area, instructional management, that Bayesian procedures for determining proficiency decisions would best seem to reside. Work in this area has been concerned with how the computer can assist in the planning and subsequent monitoring of both short- and long-term instruction for individual students. Thus, it would seem appropriate to incorporate a decision-making procedure concerned with individual proficiency level in some skill, or set of skills, as an element of the instructional management component of the IPI program. Specifically, the computer might be used to receive test data on a student and combine this with previously acquired information on other students in this IPI program, analyze the data using Bayesian analysis techniques, and then print out a report indicating the confidence which one could place in deciding that the student is proficient in a given skill at some prespecified level of performance. A more detailed discussion of how this procedure might work is now provided in the context of IPI posttests. Procedures similar to those described below would apply for placement tests and pretests as well.

Development and use of a posttest profile. The primary purpose for administering a placement test, a pretest, or a posttest is to acquire data which can be used to evaluate a student's instructional needs. When a student is administered a posttest, he is presumed to have had instruction in those skills for which he lacked sufficient proficiency at the time he was administered the unit pretest. The posttest either affirms the student's success in acquiring the skills or calls attention to those skills in which additional work is required before he can proceed to the next unit. Thus, the only information which the teacher and student need is a simple statement regarding the level of proficiency at which the student has performed on each skill in the unit. Figure 2 shows an IPI posttest profile based on a test consisting of five, eight-item subtests, each measuring proficiency level on a particular skill.

Level E-Multiplication/Division

Skill	Percent Correct
1	87.5
2	87.5
3	75.0
4	100.0
5	67.5

Fig. 2. Sample of Posttest Profile Currently in Use in IPI.

Presently, the posttest profile names each skill in the unit and lists the *percentage* of items which the student answered correctly. Given the sample profile in Figure 2 and a criterion (cutoff) score of 85%, it is likely that the student would be called upon to undertake additional work in the 3rd and 5th skills of the unit.

Under the proposed change, rather than evaluating student proficiency solely on the posttest results, additional data would be incorporated within the decision analysis process, and furthermore, the quantity reported would be an index relating the student's estimated proficiency to a stipulated standard. However, it should be emphasized that although the nature of the data reported in the student profile would change, the procedures employed by the teacher and/or student to judge proficiency would remain the same. Specifically, the posttest profile, which presently contains a statement of the percentage of items correctly answered for each skill, would be altered to report the probability that the student has achieved some prespecified level of proficiency in each objective. As far as the teacher or student is concerned, the proficiency decision process is exactly the same—judgments are based on the

evaluation of a single number or "index" for each skill Figure 3 provides an example of such a profile.

Level E-Systems of Measurement

Skill	Mastery Index
1	80
2	90
3	76
4	92
5	40

Fig. 3. Proposed Sample Posttest Profile Using Bayesian Decision Analysis Procedures.

In Figure 3, the column labeled Mastery Index actually represents a probability statement. If, for example, the criterion or cutoff score for sufficient proficiency is .85, the Mastery Index column gives the probability that the student's level of proficiency is above .85 for each skill. In this case, the mastery index for skill 1 is .80. We see that the actual test performance was only 75%. This might suggest, very roughly, a probability of .50, a 50/50 chance, for the true level of functioning being above .75. However, the Bayesian analysis, using the collateral information has raised to .80 the probability that the student's level of functioning is above .85. Therefore, if we would want to move a student on if the odds were better than three to one in favor of his actually being proficient, we would advance this student since his probability of mastery is greater than .67.

Implementation mode. A profile similar to the one described in Figure 3 could be provided in at least two ways. One method of delivery would require the availability of tests which are administered by computer Presently, test administration by computer is very much a part of the feasibility study underway in IPI Given the existence of a unit posttest on some specified unit, it would seem quite possible for sample data generated by the computer test to be merged with a file containing collateral data on student success in the system. For example, the computer test program could be designed, upon student completion of the test, to call a subroutine which would access the collateral data file, combine the two sets of information, compute the mastery indices (aposteriori probabilities), and print out a profile similar to Figure 3. In this case, the collateral data would be in a file permanently maintained on the computer and periodically updated. This function could be performed automatically by the computer.

Since it is very likely that many schools using IPI will not have ready on-line access to a computer, an alternative procedure for providing the same decision analysis would call for the construction of simple "Mastery Index" tables. These tables would permit the teacher, the aide, or a student to determine the probability that the student has sufficient proficiency in a skill by simply entering the table with the number of items answered correctly on each skill of the posttest. Figure 4 serves as an example of such a table.

Level E-Systems of Measurement

Skill 1

Number of items answered correctly	Mastery Index
8	.98
7	.93
6	.85
5	.73
4	.60
3	.34
2	.27
1	.12
0	.03

Fig. 4. Sample of Proposed "Mastery Index" Table for IPI.

Given knowledge of the number of items which the student answered correctly out of a possible eight on skill 1 of the level E posttest for Systems of Measurement, the teacher or student would enter the "Mastery Index" table with that number. For example, if the student responded correctly to seven of eight items, he would enter the table in the left hand column with the number seven and consequently determine that the probability that the student has the prespecified level of proficiency, say .85, is 93. The decision as to whether to move a student forward or not would depend on this probability and the relative disutilities associated with the two kinds of errors. The simple methods for accomplishing this are described by Davis, Hickman, and Novick (1973).

The indices reported in the tables would have been generated at some earlier time and would have included consideration of relevant prior data regarding student success on the skills contained in the unit. The tables would be updated on a regular basis as increased numbers of students proceeded through the system, thus making more prior information available. Such an updating might occur once or twice a year.

SUMMARY

Individualized learning programs like IPI generate substantial amounts of data related to student success on skills in the system. Given these data, it seems reasonable to suggest that they should be used to improve the quality of instructional decision making. In particular, prior data should be combined with sample test data to form a more complete information base on which to evaluate student proficiency. By using such data jointly, instructional decisions regarding a student's needs as they relate to a given skill or set of skills will be deserving of more confidence than present decisions which are currently based solely on the student's performance on a short test.

Two procedures for implementing such a plan have been proposed. One calls for the marriage of the Bayesian decision analysis procedures with computer administered tests, whereas, the other would rely on the teacher or student to consult a table to translate student test performance to a "Proficiency Index" which would incorporate both the test data and prior data regarding student success in the system. The ultimate criterion for success of such a plan is the extent to which it leads to improvements in the instructional decision process. To this end, the next step is to implement the procedures and evaluate their impact on students within IPI.

REFERENCES

Block, K., Carlson, M., Fitzhugh, R., et al. *A computer resource for the elementary school. Progress report.* Pittsburgh, Pa.: University of Pittsburgh, Learning Research and Development Center, 1973.

Cooley, W. W., & Glaser, R. Computer and individualized instruction. *Science*, 1969, **166**.

Davis, C., Hickman, J., & Novick, M. R. A primer of decision analysis for individually prescribed instruction. *ACT Technical Bulletin No. 17.* Iowa City, Iowa: The American College Testing Program, 1973.

Ferguson, R. *Computer assistance for individualizing measurement.* Pittsburgh, Pa.: University of Pittsburgh, Learning Research and Development Center, 1970. (a)

Ferguson, R. Computer assistance for individualizing instruction. *Computers and Automation*, March 1970, **19**. (b)

Ferguson, R. A model for computer-assisted criterion-referenced measurement. *Education*, 1971, **81**, 25-31.

Ferguson, R., & Hsu, T. C. *The application of item generators for individualizing mathematics testing and instruction.* Pittsburgh, Pa.: University of Pittsburgh, Learning Research and Development Center, 1971.

Flanagan, J. Functional education for the seventies. *Phi Delta Kappan*, September 1967.

Glaser, R. Adapting the elementary school curriculum to individual performance. In *Proceedings of the Invitational Conference on Testing Problems.* Princeton, N.J.: Educational Testing Service, 1968.

Glaser, R., & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.) *Educational Measurement.* Washington: American Council on Education, 1971.

Hambleton, R. K. A review of testing and decision-making procedures for selected individualized instructional programs. *ACT Technical Bulletin No. 15*. Iowa City, Iowa. The American College Testing Program, 1973.

Hambleton, R. K., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. *Journal of Educational Measurement*. 1973, in press.

Lewis, C., Wang, M. M., & Novick, M. R. Marginal distributions for the estimation of proportions in *m* groups. *ACT Technical Bulletin No. 13*. Iowa City, Iowa. The American College Testing Program, 1973.

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in *m* groups. *Psychometrika* 1973, **38**, 19-46.

Wang, M. M. Tables of constants for the posterior marginal estimates of proportions in *m* groups. *ACT Technical Bulletin No. 14*. Iowa City, Iowa. The American College Testing Program, 1973.

Wang, M. M., & Lewis, C. Estimation of proportions in two-way tables. *ACT Technical Bulletin No. 16*. Iowa City, Iowa. The American College Testing Program, 1973. (a)

Wang, M. M., & Lewis, C. Marginal distribution for the estimation of proportions in a two-way table. *ACT Technical Bulletin No. 19*. Iowa City, Iowa: The American College Testing Program, 1973, in press. (b)

83

ACT Research Reports

This report is Number 60 in a series published by the Research and Development Division of The American College Testing Program. The first 26 research reports have been deposited with the American Documentation Institute, ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D.C. 20540. Photocopies and 35 mm microfilms are available at cost from ADI, order by ADI Document number. Advance payment is required Make checks or money orders payable to. Chief, Photoduplication Service, Library of Congress. Beginning with Research Report No 27, the reports have been deposited with the National Auxiliary Publications Service of the American Society for Information Science (NAPS), c/o Microfiche Publications, 305 East 46th Street, New York, New York 10017. Photocopies and 35 mm microfilms are available at cost from NAPS Order by NAPS Document number. Advance payment is required. Printed copies ($1 00) may be obtained, if available, from the Publications Division, The American College Testing Program, P.O. Box 168, Iowa City, Iowa 52240. A check or money order must accompany the request.

The reports since October 1970 in this series are listed below. A complete list of the reports can be obtained by writing to the Publications Division, The American College Testing Program, P.O. Box 168, Iowa City, Iowa 52240.

No. 37 *Practices and Outcomes of Vocational-Technical Education in Technical and Community Colleges*, by T G Gartland, & J. F. Carmody (NAPS No. 01441; photo, $6.80; microfilm, $2.00)

No. 38 *Bayesian Considerations in Educational Information Systems*, by M. R. Novick (NAPS No 01442, photo, $5 00, microfilm, $2.00)

No. 39 *Interactive Effects of Achievement Orientation and Teaching Style on Academic Achievement*, by G Domino (NAPS No. 01443; photo, $5.00; microfilm, $2.00)

No. 40 *An Analysis of the Structure of Vocational Interests*, by N. S. Cole, & G. R. Hanson (NAPS No 01444, photo, $5 00, microfilm, $2.00)

No. 41 *How Do Community College Transfer and Occupational Students Differ?* by E J Brue, H B Engen, & E J Maxey (NAPS No. 01445; photo, $5.50; microfilm, $2.00)

No. 42 *Applications of Bayesian Methods to the Prediction of Educational Performance*, by M R Novick, P H Jackson, D T. Thayer, & N. S. Cole (NAPS No. 01544, photo, $5.00; microfilm, $2.00)

No. 43 *Toward More Equitable Distribution of College Student Aid Funds. Problems in Assessing Student Financial Need*, by M. D. Orwig (NAPS No. 01543; photo, $5.00; microfilm, $2.00)

No. 44 *Converting Test Data to Counseling Information*, by D. J. Prediger (NAPS No 01776, photo, $5 00, microfiche, $2.00)

No. 45 *The Accuracy of Self-Report Information Collected on the ACT Test Battery: High School Grades and Items of Nonacademic Achievement*, by E. J. Maxey, & V. J. Ormsby (NAPS No. 01777, photo, $5.00, microfiche, $2.00)

No. 46 *Correlates of Student Interest in Social Issues*, by P H. Fenske, & J. F. Carmody (NAPS No 01778, photo, $5 00, microfiche, $2.00)

No. 47 *The Impact of College on Students' Competence to Function in a Learning Society*, by M H Walizer, & R E Herriott (NAPS No. 01779; photo, $5.00; microfiche, $2.00)

No. 48 *Enrollment Projection Models for Institutional Planning*, by M. D. Orwig, P K. Jones, & O T Lenning (NAPS No 01780; photo, $5.00; microfiche, $2.00)

No. 49 *On Measuring the Vocational Interests of Women*, by N. S. Cole (NAPS No. 02071, photo, $5 00, microfiche, $1.50)

No. 50 *Stages in the Development of a Black Identity*, by W. S. Hall, R. Freedle, & W. E. Cross, Jr (NAPS No 02072, photo, $5.00; microfiche, $1.50)

No. 51 *Bias in Selection*, by N. S. Cole (NAPS No. 02073; photo, $5.00; microfiche, $1.50)

No. 52 *Changes in Goals, Plans, and Background Characteristics of College-Bound High School Students*, by J F Carmody, R. H. Fenske, & C. S. Scott (NAPS No. 02074; photo, $5.75; microfiche, $1.50)

No. 53 *Toward an Integration of Theory and Method for Criterion-Referenced Tests*, by R K Hambleton, & M R Novick (NAPS No. 02075; photo, $5.00; microfiche, $1.50)

No. 54 *College Student Migration*, by R. H. Fenske, C. S. Scott, & J. F. Carmody (NAPS No 02215, photo, $5 00, microfiche $1.50)

No 55 *Predictions of Performance in Career Education*, by M. R. Novick, P. K. Jones, & N. S. Cole (NAPS No 02216, photo, $5.00; microfiche, $1.50)

No 56 *Predictors of Graduation from College*, by E. Nicholson (NAPS No. 02217, photo, $5.00, microfiche, $1.50)

No 57 *Schooling and Subsequent Success. Influence of Ability, Background, and Formal Education*, by L. C. Solmon (NAPS No. 02218; photo, $5.00; microfiche, $1 50)

No 58 *Common Fallacies about Heredity, Environment, and Human Behavior*, by A. Anastasi (NAPS No. 02220, photo, $5.00; microfiche, $1.50)

No 59 *A Study of the College Investment Decision*, by W. W. McMahon & A. P. Wagner (NAPS No. 02219, photo, $5.00, microfiche, $1.50)

ACT TECHNICAL BULLETIN NO. 15

A REVIEW OF TESTING AND DECISION-MAKING PROCEDURES FOR SELECTED INDIVIDUALIZED INSTRUCTIONAL PROGRAMS

by

Ronald K. Hambleton
The University of Massachusetts

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

August, 1973

A Review of Testing and Decision-Making Procedures for Selected Individualized Instructional Programs[1]

Ronald K. Hambleton[2]
University of Massachusetts

I. Introduction

1.1 Background

While the idea of developing instructional programs in our schools to meet individual student needs is not a new theme in American education (see, for example, Washburne, 1922; and Wilhelms, 1962), it has only been in the last decade that such programs have been implemented on any large-scale basis in the schools.

The basic argument in favor of individualizing instruction comes from a multitude of research studies that suggest that students differ in interests, motivation, learning rate, goals, and capacity for learning among other things; and, therefore, grouped-based instruction on a common curriculum is inappropriate to meet their educational needs. That change in our schools is obvious when one notes that schools provide successful learning experiences for only about one-third of our students (Block, 1971).

---

[1]The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U.S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

[2]The author would like to acknowledge the insightful comments and constructive criticisms of Melvin R. Novick of The American College Testing Program on earlier drafts of the manuscript. In addition, Richard Ferguson and Roy Williams provided many useful suggestions.

On the basis of Project TALENT data, Flanagan, et al., (1964) reported that our current instructional programs are inadequate to handle the large individual differences in any age or grade group. In addition, schools generally fail to help the student develop a sense of responsibility for his educational, personal, and social development or to make realistic educational decisions and choices about his future.

This trend toward individualization of instruction in education has resulted in the development of a diverse collection of attractive alternative models (see, for example, Gibbons, 1970; and Heathers, 1972) that, according to their supporters, offer new approaches to student learning which . . provide almost all students with rewarding school experiences. These include: Individually Prescribed Instruction (IPI) (Glaser, 1968, 1970), Program for Learning in Accordance with Needs (PLAN) (Flanagan, 1967, 1969), Computer-Assisted Instruction (CAI) (Suppes, 1966; Atkinson, 1968; Atkinson and Wilson, 1969), Individualized Mathematics Curriculum Project (De Vault, Kriewall, Buchanan, and Quilling, 1969), and Mastery Learning (Carroll, 1963, 1970; Bloom, 1968; and Block, 1971). All of the models, as well as many others, represent significant steps forward in improving learning by individualizing instruction. They strive to actively involve the student in the learning process, allow students in the same class to be at different points in the curriculum, and permit the teacher to give more individual attention.

In important aspects of these individualized instructional programs such as the construction of instructional materials (Popham, 1969; Smith, 1969), curriculum design (Wittrock and Wiley, 1970) and computer management (Baker, 1971; Cooley and Glaser, 1969), there are substantial bodies of knowledge. It is perhaps surprising to note then that the

amount of information currently available on the testing methods and decision procedures for these programs is quite limited. It is this component that, in principle, facilitates the efficient movement of students through the instructional program.

One reason for a lack of information is that measurement requirements within the context of many of the new programs require new kinds of tests. These are the criterion-referenced tests which are constructed and interpreted in ways quite different from the norm-referenced tests which are more familiar to most practitioners in the field (Popham and Husek, 1969; Glaser and Nitko, 1971; Hambleton and Novick, 1973).

Since one of the major purposes of individualized programs is to maximize the opportunity for all students to learn, it follows that tests used to monitor student progress should be keyed to the instruction. Further, they should provide information that can be used to measure progress along an absolute ability continuum. Norm-referenced tests are constructed specifically to facilitate making comparisons among students; hence, they are not very well suited for making most of the instructional decisions required in individualized instructional programs.

1.2 Criterion-Referenced Testing and Measurement

Much of the discussion in the area of criterion-referenced testing and measurement (for example, see Block, 1971; Ebel, 1971; Glaser and Nitko, 1971; and Hambleton and Novick, 1973) stems from different understandings as to the basic purpose of testing in the instructional models described in the previous section. It would seem that in most cases the pertinent question is whether or not the individual has attained some prescribed degree of competence on an instructional performance task. Questions of precise achievement levels and comparisons among individuals

on these levels seem to be largely irrelevant. In many of the new instructional models, tests are used to determine on which instructional objectives an examinee has met the acceptable performance level standard set by the model designer. This test information is usually used immediately to evaluate the student's mastery of the instructional objectives covered in the test, so as to appropriately locate him for his next instruction (Glaser and Nitko, 1971). Tests especially designed for this particular purpose have come to be known as <u>criterion-referenced</u> tests. Criterion-referenced tests are specifically designed to meet the measurement needs of the new instructional models. In contrast, the better known <u>norm-referenced</u> tests are principally designed to produce test scores suitable for ranking individuals on the ability measured by the test. A very flexible definition of a criterion-referenced test has been proposed by Glaser and Nitko (1971): "...[a test] that is deliberately constructed so as to yield measurements that are directly interpretable in terms of specified performance standards." According to Glaser and Nitko (1971), "The performance standards are usually specified by defining some domain of tasks that the student should perform. Representative samples of tasks from this domain are organized into a test. Measurements are taken and are used to make a statement about the performance of each individual relative to that domain." Distinctions between norm-referenced tests and criterion-referenced tests have been presented by Glaser (1963), Glaser and Nitko (1971), Livingston (1972), Popham and Husek (1969), Ebel (1971), Block (1971), Hambleton and Gorth (1971), and Hieronymous (1972).

Hambleton and Novick (1973) have discussed the evaluation of criterion-referenced tests in practical situations. In their formulation, reliability takes the form of an index indicating the consistency of decision making

across parallel forms of the criterion-referenced test or across repeated measurements. Validity takes the same form except, of course, that a new test or some other appropriate measure serves as the criterion. Both reliability and validity concepts are reformulated in straightforward decision-theoretic terms. However, at this stage of the development of a theory of criterion-referenced measurement, the establishment of cut-off scores is primarily a value judgment. [Further clarification is provided by Hambleton and Novick (1973), Millman (1973), and Block (1972).]

1.3 Instructional Models Under Consideration

The major concern in this paper is with instructional models that include a specification of the curriculum in terms of behavioral objectives, detailed diagnosis of the entering competencies of students, the availability of multiple instructional resources, individual pacing and sequencing of material, as well as the careful monitoring of student progress.

In the programs under consideration, Computer-Managed Instruction (CMI) is an optional feature. Under CMI the goal is for the computer to service classroom terminals which assist the classroom teacher in assessing a student's strengths and weaknesses, and to prescribe instructional sequences (Cooley and Glaser, 1969). Project PLAN and CAI are implemented in a CMI mode whereas IPI and Mastery Learning are not.

In summary, the goals of individualized instructional programs developed along the general lines of the specifications above are to enable students to work through the units of instruction at a pace

reasonable for them, to develop self-direction and self-initiation, to encourage self-evaluation as well as motivation for learning, and to demonstrate mastery in a variety of skills.

Cronbach (1967) reported on three major patterns of dealing with individual differences which provide a framework for the models considered in this paper. Patterns of dealing with individual differences in the school can be described in terms of the extent to which educational goals and instructional methods are varied. In one pattern, the educational goals and instructional methods are relatively fixed and inflexible. Individual differences are handled mainly by dropping students from the program when they begin to encounter difficulty. In a second pattern, goals are selected for students on the basis of interest and potential. They are then channeled into one fixed program or another. Individual differences are handled by providing multiple optional programs. The models we describe in this paper fit into a third pattern where goals <u>and</u> instructional resources are individualized for the purpose of maximizing learning.

1.4 Purposes of the Investigation

The success of individualization depends to a considerable extent on how effectively teachers and students make decisions as to the mastery of specific instructional objectives, the development of individual prescriptions, the selection of instructional resources, etc. However, various writers including Baker (1971) and Glaser and Nitko (1971) have commented rather critically on existing testing techniques and procedures. Relevant background for improving such a situation would certainly include a review of the testing models of some of the more commonly used individualized instructional programs. Such a review would assist in

defining the kinds of decisions that are made, and the information on which the decisions are based. This should provide a basis for developing testing methods and decision procedures specifically designed for use within the context of these models. (Although it would be ideal to develop a general measurement model to cover all the instructional models, we are not prepared in this paper to advance such a model.)

The first purpose of the investigation was to provide a description of the testing models that are currently being used in selected individualized instructional programs. Three programs were selected for study: Individually Prescribed Instruction, Program for Learning in Accordance with Needs, and Mastery Learning. [These models as well as others are also discussed by Baker (1971); however, he was concerned with their computer-based instructional management systems which are of only secondary interest in this paper.] These programs were selected in this study because they are among the best known and because there is a substantial amount of information available on each. In the following sections, an introduction is provided for each instructional model. The introduction includes a brief history of the program, the content areas covered, and an indication of the extent of implementation. Also, a description of each instructional paradigm and details on the testing model is provided. An attempt is made to pinpoint the decision points in each model, spelling out the consequences of the various possible actions in relation to each of the "possible true states of nature."

The discussion of the models is based on descriptions found in books, papers, and reports; on-site visits; and meetings with many of the developers. It should be noted however that programs are often implemented by teachers quite differently than they are reported in

the literature. Also, it should be remembered that these programs are constantly changing; hence, it is possible that certain features of the models are not exactly as they are described here. In particular, it is our impression that PLAN is being implemented in a way quite different from how it has been described in the literature. This is because Westinghouse Learning Corporation has now taken over the development and implementation components.

A second purpose was to compare the three programs and the four component parts of the testing model; namely, selection of a program of study, criterion-referenced testing on the unit objectives, assignment of instructional modes, and final year-end assessment.

A final purpose was to briefly outline several promising lines of research in connection with the testing methods and decision procedures for individualized instructional programs.

II. Individually Prescribed Instruction (IPI)

2.1 Background

The Learning Research and Development Center (LRDC) at the University of Pittsburgh initiated the Individually Prescribed Instruction Project during the early 1960's at the Oakleaf School in cooperation with the Baldwin-Whitehall Public School District near Pittsburgh. Major contributors to the project over the years include Robert Glaser, John Bolvin, C. M. Lindvall, and Richard Cox. Initial activities concentrated on producing instructional materials and training materials. More recently, research and evaluation activities have assumed an increasingly important role in Center activities.

As of 1972 the IPI program was being implemented in over 250 schools around the country. Distribution of materials and other information on the program is managed by Research for Better Schools, Inc., a United States Office of Education Regional Laboratory located in Philadelphia. At present, instructional materials are available in elementary mathematics, reading, science, handwriting, and spelling.

2.2 Description of the Instructional Paradigm

While we will discuss the instructional paradigm and the corresponding test model in the context of the IPI mathematics program, the procedures, techniques, etc., described, are in no way limited to that content area. In fact, it should be noted that the mathematics program as implemented is probably somewhat different from what we describe here, since the LRDC is constantly refining and improving the program (Lindvall, personal communication). Fortunately, for our purposes the basic structure of the program remains as described.

It is instructive first of all to describe the structure of the mathematics curriculum. Cooley and Glaser (1969) report that the mathematics curriculum consists of 430 specified instructional objectives. These objectives are grouped into 88 units. (In the 1972 version of the program there were 359 objectives organized into 71 units.) Each unit is an instructional entity which the student works through at any one time. There are 5 objectives per unit, on the average, the range being 1 to 14. A collection of units covering different subject areas in mathematics comprises a level; the levels may be thought of as roughly comparable to school grades. For illustrative purposes, Table 2.2.1 presents the number of objectives for each unit in the IPI mathematics curriculum.

The teacher is faced with the problem of locating for each student, that point in the curriculum where he can most profitably begin instruction. Also, the teacher is responsible for the continuous diagnosis of pupil demonstrating proficiency in each skill prescribed in his particular instructional sequence as he moves along.

At the beginning of each school year the teacher places the student within the curriculum; that is, he identifies the units in each content area for which instruction is required. After completing the gross placement, a single unit is selected as the starting point for instruction, and a diagnostic instrument administered to assess the student's competencies on objectives within the unit. The outcome of the unit test is information appropriate for prescribing instruction on each objective in the unit. In addition it is also necessary to select the particular set of resources for the student. In theory, resources that match the individual's "learning style" are selected. Within each unit, there are short tests to monitor the student's progress. Finally, upon completion of initial in-

Table 2.2.1[1]

Number of Objectives for Each Unit in the IPI Mathematics Curriculum

Content Area	Levels							
	A	B	C	D	E	F	G	H
Numeration	12	10	8	8	8	3	8	4
Place Value		3	5	10	7	5	2	1
Addition	3	10	5	8	6	2	3	2
Subtraction			4	6	3	1	3	1
Multiplication				8	11	10	6	3
Division				7	7	9	5	5
Combination of Processes			6	5	7	4	5	6
Fractions	3	2	4	6	6	14	5	2
Money		4	4	6	4	1		
Time		3	2	7	9	5	3	1
Systems of Measurement		4	3	5	7	3	2	
Geometry		2	2	3	9	10	7	9
Special Topics			1	3	3	5	4	5

[1] Reproduced, by permission, from Lindvall, Cox, and Bolvin (1970).

struction in each unit, assessment and diagnostic testing takes place. In the next section, we review the tests and the mechanisms for making these decisions. Suffice to say here that it has been found that teachers differ in the extent to which they follow prescribed decision-making rules (Lindvall, Cox, and Bolvin, 1970).

2.3 Details of the Testing Model

Various reports over the last couple of years have dealt with the testing model and its development (Lindvall, Cox, and Bolvin, 1970; Glaser and Nitko, 1971; Cox and Boston, 1967). A flow chart of the testing model is presented in Figure 2.3.1. To monitor a student through the program the following tests are used: placement tests, unit pretests, unit posttests, and curriculum-embedded tests. All of the tests are criterion-referenced with performance on the tests compared to performance standards for decision-making.

How sophisticated is the decision-making process utilizing the scores from the various tests? According to Glaser (1968):

> At the present stage of our knowledge, the decision rules for going from measures of student performance to instructional prescriptions may not be very complex, but little is known about the amount of complexity required, although, the individual monitoring of student performance provides us with a good data base to study this process.

Promising developments in the last couple of years include increased knowledge about constructing and evaluating criterion-referenced tests. Also, the research on branched testing strategies (Ferguson, 1969, 1971) has much potential for improving the efficiency of the testing model. This second point will be discussed in greater detail in a later section.

Placement Tests

When a new student enters the program, it is necessary to place the student at the appropriate level of instruction in each of the content areas.

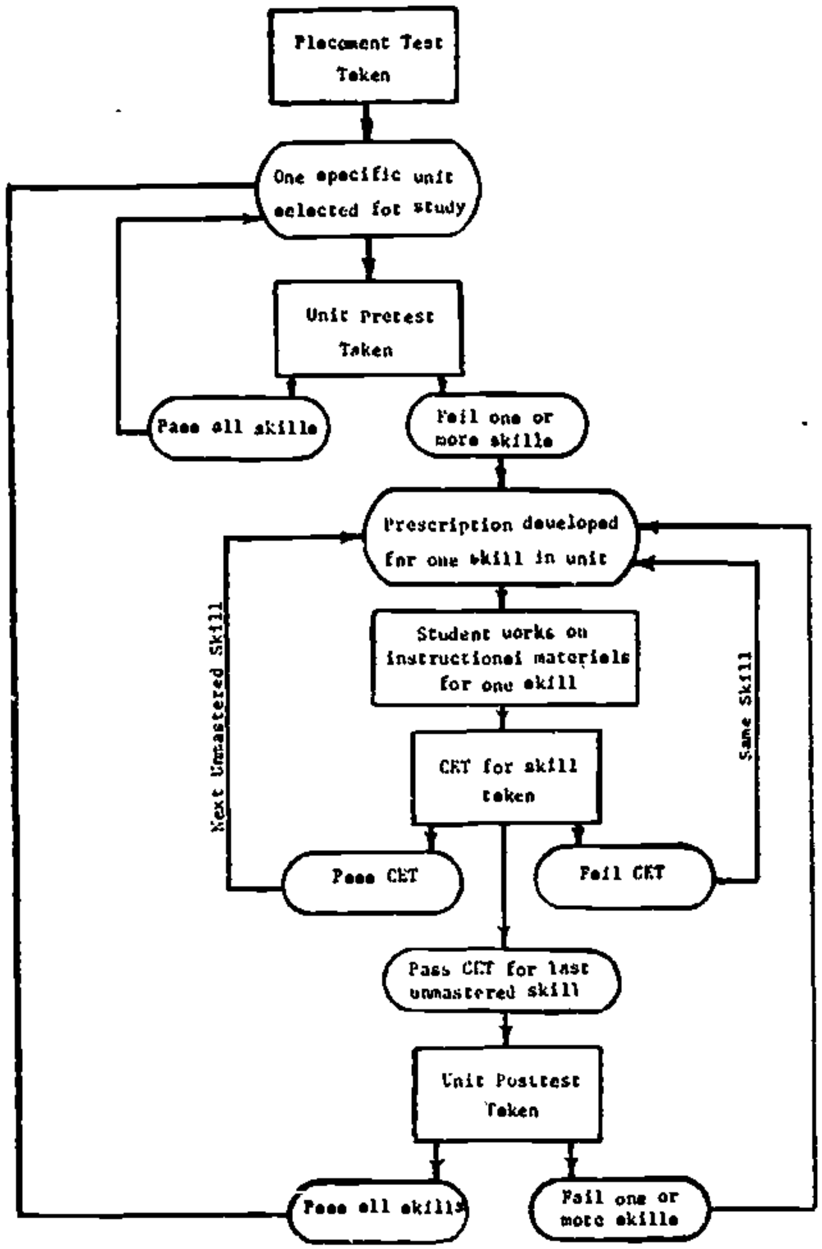


Figure 2.3.1. Flow chart of steps in monitoring student progress in the IPI program. (Reproduced, by permission, from Lindvall and Cox, 1969.)

[Glaser and Nitko (1971) called this stage-one placement testing.] Typically, this is done by administering a placement test which covers all of the subject areas at a particular level (see Table 2.2.1). Factors affecting the selection of a level for placement testing of a student include student age, past performance, and teacher judgment. Generally, the placement test covers the most difficult or most characteristic objectives within each area. Placement tests are administered until a unit profile identifying a student's competencies within each area is complete. At present, the somewhat arbitrary 80-85% proficiency level is used for most tests in the IPI system.

Scores for a student on items measuring objectives in each unit and area in the placement test are used to define an individual program for him. The standard procedure is to assign instruction on units in which placement test performance on items measuring a few representative objectives in the units is between 20% and 80%. If the score is less than 20% for a given unit, the unit test in the area at the next lowest level is administered and the same criterion is applied. If he passes the unit test, he receives instruction in the unit in the next level. In the case where a student has a score of 80% or over, he is tested on the unit in the area at the next highest level. [Further information is provided by Lindvall, Cox, and Bolvin (1970), Weisgerber (1971) and Cox and Boston (1967).]

For example, suppose a student were to achieve scores on level E of 60%, 90%, 60%, 60%, 30%, 30%, 25%, 90%, 50%, 10%, 0%, 30%, 30% in the thirteen areas indicated in Table 2.2.1. It is likely that he would be prescribed instruction at level E in the areas of numeration, addition, subtraction, multiplication, division, combination of processes, money, geometry, and special topics. He would receive the level F placement tests in place value and fractions. If, for example, he scores 60% and 10% respectively, he would receive instruction at level F in place value and probably at

level E in fractions. He would also be administered the level D placement tests in the areas of time and systems of measurement. If, for example, his scores were 0% and 40%, he would receive a still lower placement test in the area of time and would be prescribed instruction at level D in systems of measurement. If he scores 85% on the level C placement test in the area of time, he would be assigned to level D for instruction.

In order to acquire some information on the average length of the tests, the level E placement tests of the 1972 edition of the IPI program were selected and examined. Analysis revealed that on the average there are 12 items measuring the objectives in each area (with a range of from six to 20).

In summary, we note that the placement test has the following characteristics: provides a gross level of achievement for any student in the curriculum, and provides information for proper placement of students in the curriculum.

Unit Pretests and Posttests

Having received an initial prescription of units, a student proceeds by taking a pretest for a unit at the lowest level of mastery on his profile. [Glaser and Nitko (1971) call this stage-two placement testing.] A unit pretest includes one or more items to measure each objective in the unit. A review of the unit pretests and posttests in level E revealed that the approximate number of items on a test is 37 (the range is from 21 to 64) and the average number of items measuring each objective is six (the range is from four to seven). Lindvall and Cox (1969) report that the length of a pretest is determined by the number of objectives in the instructional unit and by the number of items used to test each objective. No fixed number of items to measure each objective is used because of the diverse nature of the objectives. For example, they note that, "an objective like--the pupil can solve simple addition problems involving all number combinations--will

require more items than would an objective like--the pupil must select which of three triangles is equilateral--."

A student is prescribed instruction in each objective in the unit for which he fails to achieve an 85% mastery level.[1] In the case where the student demonstrates mastery of each objective, he is moved on to the next unit in his profile, where he again takes a pretest.

The unit posttests are simply alternate forms of the unit pretests and are administered to students as they complete instruction on the unit. A student receives a mastery score for each objective in the unit. He is required to repeat instruction on any objective where he fails to achieve an 85% mastery score. He is directed to the next unit in his profile if he demonstrates mastery on each objective covered in the unit posttest. Those who repeat instruction on one or more of the objectives must take the unit posttest again before moving on in their program.

In summary, pretests and posttests are available for each unit of instruction. The proper pretest is administered on the basis of student's curriculum profile, and learning tasks for each skill are assigned (or not assigned) on the basis of a student's performance on items measuring the skill.

Compared with students in many other types of mathematics programs, it is clear that the student in the IPI program spends more of his time taking tests. However, to some extent this can be justified on the grounds that testing is an integral part of the learning process in the IPI model. Nevertheless, there seems to be good reason for researching techniques to reduce testing time.

---

[1] A mastery score on each objective for a student is calculated as the percentage of items on the test that measure the objective that the student answers correctly.

Hsu and Carlson (1972) point out several problems associated with the current version of the unit pretests and posttests. The existing system requires that every objective be tested; hence, the time a student spends taking tests is considerable. Also, because of management and scoring problems, feedback to the student on his results is not immediate. Further, students are occasionally required to take the same posttest on a second occasion. This raises a question about practice effect.

One very promising way to reduce the testing time with the correlated result of producing better instructional decisions is suggested in the branched testing work of Ferguson (1969, 1971). Ferguson showed that by using a tailored testing strategy, a computer terminal to monitor the selection of test items, and information on the hierarchical structure of the items, he was able to significantly reduce unit testing time without any loss in decision-making accuracy. A comprehensive review of the work in branched testing is out of place here; suffice to say here that major contributions to the area include Ferguson (1969, 1971), and Lord (1970). A review of some of the work in the area is provided by Bock and Wood (1972).

Curriculum-Embedded Tests

As the student proceeds through a unit of instruction, his progress must be monitored. This is done by curriculum-embedded tests (CET). As used in the mathematics IPI program, a CET is primarily a measure of performance on one specific objective. There are usually several test items to measure the objective. A review of the CETs in level E of the program revealed that there are on the average about three items measuring the primary objective covered in the CET. The range is from two to five. If a student

receives a score of 85%, he is permitted to move on to the next prescribed objective. Otherwise, he is sent back for additional work and then he takes an alternate form of the CET when he is ready.

A secondary purpose of the CET is to pretest, in a rough way, the next objective in the learning sequence. (Objectives in a unit are arranged into a learning sequence.) Students may pretest out of the next skill in the sequence by achieving 85% or higher on the short test which makes up the second part of the CET and on part one of the CET for that skill. It would appear from a review of level E tests that there are about two items measuring the secondary objective. In cases where a student does not need instruction on the next skill, he can skip part two of the CET and move on to the part two of the CET that tests the next skill he needs for his program. This additional pretesting of an objective in the CET gives students a chance to demonstrate mastery of new skills not specifically covered in the instruction to that point and to eliminate that instruction from his program.

Student Diagnosis

Once the student has been assigned to a unit of instruction and the objectives for which he needs instruction have been identified by the unit pretest, there still remains the problem of deciding which of several instructional methods is "optimal" for him. That is, of the available instructional methods for a particular instructional unit, in which of them would a student with a known background in the program and specific goals, interests, and aptitudes stand the "best" chance of learning the material? Glaser and Nitko (1971) call this a diagnostic decision.

III. Program for Learning in Accordance with Needs (PLAN)

3.1 Background

Project PLAN is a major ungraded, computer-supported individualized instruction program in education developed by the American Institutes for Research over the last seven years. (For background, see Weisgerber, 1971.) The project was initiated by John Flanagan to handle many of the shortcomings of our educational system as revealed by Project TALENT (Flanagan, et al., 1964).

The PLAN program is currently being used in over 70 schools with more than 35,000 students in grades one through twelve. Instructional materials are available in four areas: social studies, language arts, mathematics, and science. Westinghouse Learning Corporation is now responsible for the monitoring and marketing of Project PLAN materials. They also operate the computer installation necessary for the proper functioning of Project PLAN in a school.

Unfortunately, the implementation of the model in 1972-73 involves far fewer features than was originally described by the proponents of the program a few years ago. Nevertheless, we will describe the more elaborate version of the program in this paper.

3.2 Instructional Paradigm

The basic unit of instruction in PLAN, called a module, is an instructional package made up of about five behavioral objectives. It normally takes a student about two weeks to complete a module of instruction. Also, there are many objectives classified at the higher levels of Bloom's (1956) taxonomy that do not fit nicely into the regular modules. These are

named module-set objectives, and examples include concept development and problem-solving skills. They are worked into the regular modules and progress is measured by PLAN achievement tests administered periodically throughout the program. According to Rhetts (1970) there are more than 1100 modules in PLAN. For each module, there are several different teacher-learning units (TLU) assigned individually on the basis of aptitudes, interests, learning style, etc. All modules in the secondary school curricula are coded as to whether, 1) they are part of a state or local requirement, 2) essential for a given educational or occupational area, 3) highly desirable for that area, 4) essential for minimum functioning as a citizen, 5) highly desirable for all citizens to know, or 6) would make the student a particularly well informed citizen.

TLU's are coded according to: 1) reading difficulty, 2) degree to which it requires teacher supervision, 3) its media richness, 4) degree to which it requires social involvement and/or group learning activities, 5) the amount of reading involved, and 6) variety of activities in the module. There are, on the average, two TLU's for each module. Along the lines of Dunn (1970), we will describe the most complex version of the program--the version currently being used in the secondary school.

At the beginning of each year, a program of study is prepared for each student. This includes a list of modules, suggested TLU's, and a recommended sequence in the four content areas. To really provide individualized instruction, it is necessary to know about student needs, goals, abilities, and interests <u>and</u> to use the information in developing a program of study (POS) for him. As part of the PLAN system then, the following information is collected:

1. parent and student educational goals
2. parent and student vocational aspirations
3. student level of achievement and vocational interests
4. student abilities (such as reading comprehension and arithmetic reasoning)
5. past performance of student in program
6. student's learning style.

A variety of questionnaires and testing instruments have been developed to collect the above information.

Abilities are measured each year with the Developed Abilities Performance Test (DAPT). This test consists of 18 scales (see, for example, Jung, 1970) such as those to measure arithmetic reasoning, reading comprehension, abstract reasoning, mechanical comprehension, and ingenuity.

On the basis of the above information, a program is developed and the student is monitored through it by continuous module posttesting and PLAN achievement testing. Let us look now at the testing phase of the program in more detail.

3.3 Testing Model Details

Within a PLAN school, there exists a multitude of decisions to make on each student. These include development of a program of study, periodic assessment of module-set objectives, performance on the modules of instruction, assignment of TLU's, and yearly monitoring of important skills. The major decision points are shown in Figure 3.3.1. Unfortunately, there is little available information on how these decisions are made.

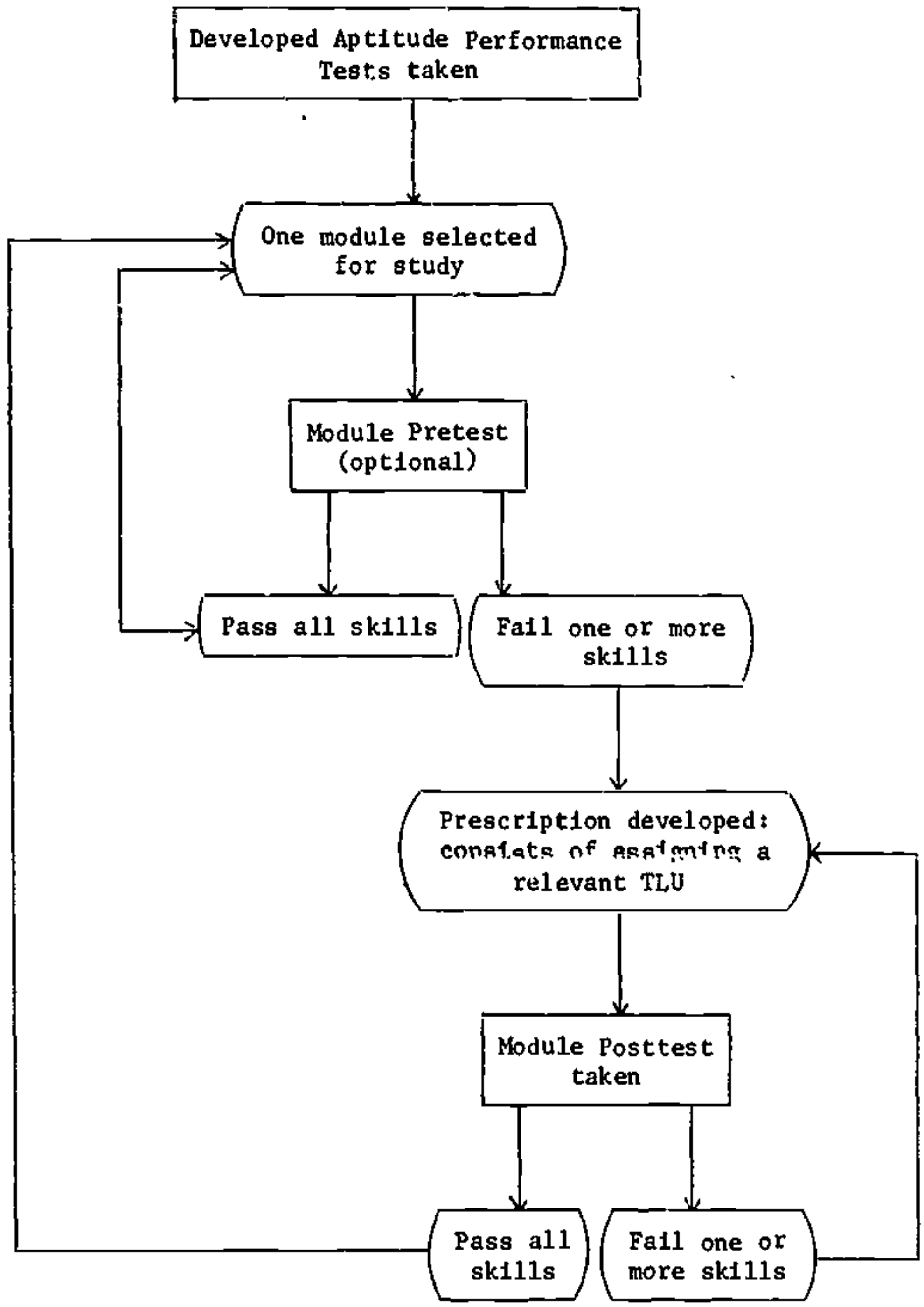


Figure 3.3.1 Flow chart of steps in monitoring student progress in Project PLAN.

Development of a Program of Study

On the basis of DAPT scores which are matched to Project TALENT data of people in different occupations, the students and parents select a long range goal [(LRG) (one of 12 families of occupations)]. Information on the long range goal along with parent and student information described in the last section is used to develop a program of study. The DAPT is also used in the determination of the number of modules a student will study in a year. Jung (1970) reports that on the basis of weights derived from regression analyses, a quota is identified for each PLAN student in each subject area. Modules are then assigned to him on the basis of his LRG group membership until this quota is filled.

Developed Aptitude Performance Tests

These tests are given at the beginning of each school year. Information on the length, kinds of test items, reliability and validity does not appear to have been published. Also, we do not know whether a different version of the test is used in each year, or whether the same version is used for several years. Regardless, unless comparability of the score scales for the different versions has been carefully done, we doubt whether the change scores (for individuals or groups) on each variable from year to year have very much meaning.

PLAN Achievement Tests

Mastery of the module-set objectives is measured at specific points in the curriculum using PLAN achievement tests. However, we are also unclear on the make-up of the PLAN achievement tests. Apparently, they are measured at "specified points" in the curriculum and the format of these tests is sometimes something other than the paper and pencil variety.

Module Tests

When the student feels he has mastered the materials covered in a module, he can take a criterion-referenced module posttest which has on it several items measuring each objective in the module. The items are presented usually in a selection format to facilitate computer scoring. On the basis of his performance, the computer using built-in decision rules makes one of four decisions. If he answers all items correctly, he is given a "complete" on the module and the computer print out tells him where to go next. If he makes a "few" errors, he is given a result of "Student Review". The computer specifies his performance on each objective and indicates the ones he should review before beginning his next module.

Students who miss a large number of items on the test but still score high enough to pass, receive a result of "Teacher Certify". He is instructed by the teacher on which objectives to review and/or restudy. He is not given his next module until, in the judgment of the teacher, he has mastered all of the objectives. An alternative is to have the student repeat the module posttest. The fourth possibility is student failure to pass the test. In this situation, he is instructed to restudy the module with the same TLU or another. In the case where he misses the test again, the teacher intervenes and takes some appropriate action to clear up the problem.

Assignment to Instructional Modes

The basic problem was described in a discussion of the IPI program, i.e., what particular instructional mode (or in this case, TLU), should the student take to study the module so as to maximize his changes of learning the material? Dunn (1970) notes, "that the computer, from a complex set of decision rules, matches the student with specific TLU's". We wonder what those rules would be, particularly since there is no theory of instruction to guide in developing optimal assignment rules. To this point in time

educational psychologists have only been able to find a handful of interactions between background variables and instructional method. A partial answer is provided by Weisgerber and Rahmlow (1971). They noted that teacher-learning units are based upon different assumed learning styles of students and are guided by a philosophy of education (Flanagan, 1970) and a theory of learning (Gagné, 1965).

IV. Mastery Learning

4.1 Background

The mastery learning concept was introduced to American Schools in the the 1920's with the work of Washburne (1922). However, because technology was not developed to the point that the program could operate efficiently, interest in the concept steadily diminished until it was revived in the form of programmed instruction in the late 1950's. (Programmed instruction was an attempt to provide students with instructional materials that would allow them to move at their own pace and receive constant feedback on their level of mastery.) The work by Carroll (1963, 1970) and Bloom (1968) and Bloom's students (Block, 1971; Airasian, 1971 and others) was instrumental in bringing mastery learning to the attention of instructional designers and researchers.

Since Bloom's paper in 1968, a great deal of research has been conducted; and the results suggest that the mastery learning model "can be easily and inexpensively implemented at all levels of education and in subjects ranging from arithmetic to philosophy to physics (Block, 1970). The model has been used now with more than 20,000 students.

4.2 Instructional Paradigm

This model is quite different from IPI and PLAN in that it attempts to individualize instruction within a group-based instructional environment. The curriculum is organized into units of instruction defined by homogeneous clusters of behavioral objectives. For each unit one or more criterion-referenced tests is used to measure mastery. Individualization is handled via supplemental materials, feedback, and corrective techniques applied to students who do poorly on the posttests.

Mayo (1970) in describing the mastery learning model notes that:

1. Students are made aware of course and unit expectations, so that they view learning as a cooperative rather than as a competitive venture.
2. Standards of mastery are set in advance for the students, and grading is in terms of absolute performance rather than relative performance.
3. Short diagnostic tests are used at the end of each instructional unit.
4. Additional learning is prescribed for those who do not demonstrate unit mastery.
5. Additional time for learning is prescribed to students who seem to need it.

The mastery learning model is less impressive in scope than PLAN, and the requirements for an effective testing plan are less stringent than with IPI or PLAN. Features of mastery learning appear to be that it is easily implementable, does not require the use of a computer, and is appropriate for almost any content area. Also, if mastery learning is carried out properly, previous research suggests that students will achieve higher scores and have more interest in school and a better attitude toward school. Unlike the other two models, with mastery learning much of the work has been on research related to the correctness of the model of school learning. An extensive number of content areas have been studied.

It should be noted that there are many variations on the basic mastery model as originally proposed by Bloom (1968). Some of them are summarized by Block (1971), and an example would be the work of Kim (1971).

4.3 Test Model Details

Block (1971) notes that, "To individualize instruction within the context of ordinary group-based instruction, mastery learning relies heavily on the constant flow of feedback information to teacher and learner." It does not seem however that there is as much testing in mastery learning as in IPI or PLAN. A flow chart of the testing component is shown in Figure 4.3.1.

The mastery learning testing model as described by Airasian (1971) represents a special case of the IPI testing program. There is no placement testing, and unit pretesting and curriculum-embedded testing are not emphasized. Unit posttesting and final assessment represent the two major kinds of testing in the program. In the spirit of Scriven (1967), these two areas are known as formative and summative tests. It should be noted, however, that formative tests or unit posttests, as they are called in IPI, are not used for grading. They are used for diagnosing learning difficulties only.

Formative Tests

A formative test is designed to cover the objectives over a short unit of instruction in the mastery learning program. It is used to determine whether or not a student has mastered the material and to serve as a basis for prescribing supplemental work in areas where the student is weak (Airasian, 1971). Implementers of the mastery learning model have set the passing standard anywhere from 75% to 100%. There is no set number of items or format suggested to measure each objective; however, there is a suggestion that instructional decisions are made on the basis of responses to individual items.

The formative tests in mastery learning represent the key to individualizing instruction since it is on the basis of these scores that

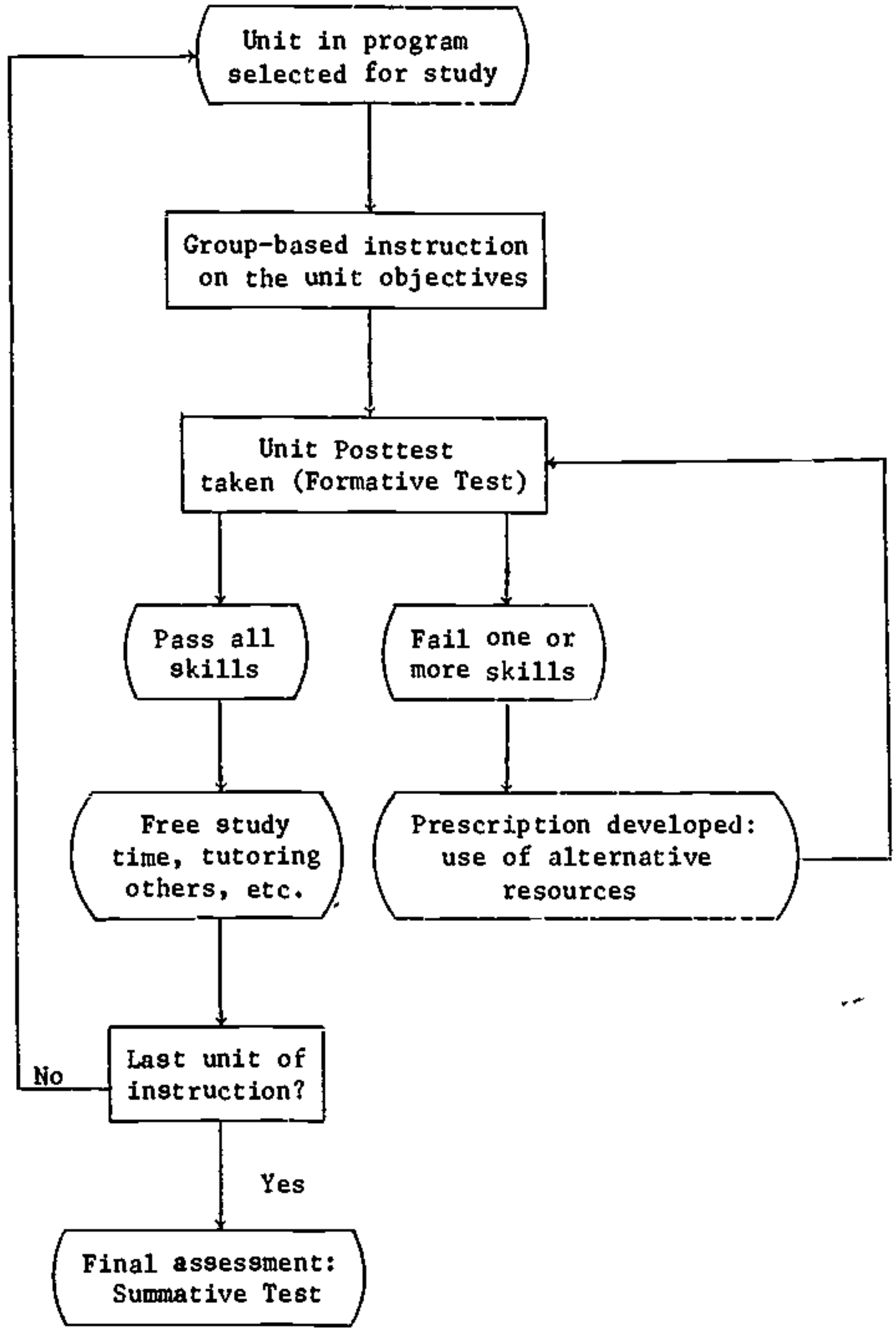


Figure 4.3.1 Flow chart of steps in monitoring student progress in a typical version of a mastery learning model.

individualization of instruction can take place. Units are kept small so that unit testing takes place frequently to increase the effectiveness of the individualization of instruction component of the program.

Summative Tests

The primary purpose of the summative test in the mastery learning model is to grade students on the basis of their achievement of course objectives. The items in the test are keyed to objectives and representative of the pool of course objectives. A criterion-referenced interpretation of the scores is recommended. It is proposed that cutting points be located on the ability continuum and grades should be assigned on the basis of a student's position on the continuum and not relative to other students in the course. A norm-referenced interpretation of the scores is also possible.

Final Comments

Mastery learning is probably the least different from traditional instructior since the principal instruction is always group-based and final grades are assigned. (However, it is expected that because of various features built into the program that the final assessment testing will not be as threatening a situation for the student as it is in more traditional programs.) Differences with traditional instructional models include features such as individual pacing, and the big difference is the use of frequency tests on small units of instruction to diagnose learning problems. Important features are the feedback/correcting-review techniques. It would appear, however, that there is little in the way of sophistication concerning the testing model. For example, there appears to be no guidelines for determining the optimum number of items to measure each objective on a unit posttest. An exception is the excellent work of

Block (1970) in investigating, among other things, the problem of setting cutting scores on criterion-referenced tests to separate students into two groups--masters and non-masters. His results suggest that setting cutting scores high (95%) may be best for cognitive learning but in the long run positive attitudes and interest in the subject are less likely to develop. With a reduction in the cutting score to 85% there was a reduction in cognitive learning, but selected affective outcomes were muximized.

V. A Comparison of the Testing Models

5.1 Introduction

In the three previous sections we have highlighted the basic testing and decision-making features in three individualized instructional programs--IPI, PLAN, and Mastery Learning. Within all three models, instruction is self-paced although mastery learning is somewhat more structured since the initial instruction on a unit is group-paced. With each of the models, the content is organized into units or modules. Generally, in IPI and ML the student is expected to demonstrate mastery on all the units before completing the program of study although by his performance on unit pretests, it is possible for him to avoid instruction on any of the units. (One variation that does come up is the availability of "enrichment materials" which are an optional part of the curriculum.) In PLAN, at any grade level there are far more units than any student could or would ever want to master. Thus, it is first of all necessary to define a content domain of study for each student.

In the remainder of the section, we shall limit discussion to testing and decision-making issues. In order to develop a framework for the discussion, we have chosen to focus on the following issues:

1) selection of a program of study;
2) criterion-referenced testing on the unit objectives;
3) assignment of instructional modes;
4) final year-end assessment.

These represent the extent of the decision paradigms within the three models. The importance and sophistication used in handling each component varies from one model to another.

5.2 A Compendium of Decision Paradigms

Selection of a Program of Study

A program of study is that collection of units which a curriculum designer deems necessary for the appropriate education of the student.

All three models are designed for utilization with a curriculum defined in terms of behavioral objectives arranged into blocks, units, or modules around a common topic or theme. Generally in IPI and ML, students are expected to demonstrate mastery in all of the available course objectives. The available course objectives define the program of study for the student. However, on the basis of high pretest results students may avoid instruction of selected units of instruction.

In PLAN, each student receives a unique program of study. The more advanced the students the more varied their programs of study become.

For reasons described above, selecting a program of study for a student in IPI or Mastery Learning is relatively easy. The decisions to be made reduce, basically, to determining whether students have mastered particular objectives. They will receive instruction only on course objectives they have not mastered. In IPI, placement tests are used to determine the level of instruction in each area for the students. Here the error of giving the student credit for units he has not mastered (a false-positive error) seems to be somewhat more serious than mistakenly assigning him to instruction he does not need (a false-negative error). This follows since a student has a second chance to demonstrate mastery of the objectives in a unit through the unit pretest if he is mistakenly assigned to study a unit he has already mastered. On the other hand, to incorrectly assign credit for mastering a unit to a student, particularly if it is an important unit, will plague him in his future studies.

In theory at least in the PLAN program, developing a program of study is a complex affair. Done once a year it requires a wealth of information described in section 3.3 to develop the program. The danger of locating a student in the wrong program because of misjudgment on the part of the parents, teachers, or the student or because of a "less than 100% prediction system" are great; however, this is the same risk we take with selection of a program in a traditional school. This is particularly serious in the high school where there is more choice than in the elementary school programs. However, the flexibility of the PLAN program makes switching from one program to another easier.

<u>Criterion-Referenced Testing on the Unit Objectives</u>

There are three kinds of testing appropriate here: unit pretesting, unit posttesting, and curriculum-embedded testing. All three kinds of testing are used in IPI and PLAN although unit pretesting is not stressed in PLAN. The possibility existed for all three kinds of testing in Mastery Learning; however, unit pretesting is not emphasized and a student can avoid the curriculum-embedded testing by passing the unit posttest and thus avoid the remedial instructional materials. (Also, it is quite possible that curriculum-embedded tests are not available in the remedial materials.)

Let us briefly look now at the losses involved in making different kinds of decisions. It should be recalled that the unit tests (or module tests) measure performance on each objective or skill with several items. On the unit pretests, a student receiving credit for non-mastered objectives will likely be "caught" on the administration of the posttest and correct instruction can be assigned at that time. However, to the extent that these objectives are prerequisites to others in the unit we have a case of instructional mismanagement. (Perhaps, this is a place where Bayesian

statistics might be helpful in producing an "improved" profile of scores across objectives measured by the unit pretest. This would undoubtedly improve the overall decision-making accuracy. Likewise this strategy could be used on the unit posttests.)

To assign a student instruction on the basis of pretest score results to objectives which he has previously mastered will undoubtedly prove to be frustrating to him; however, it should be noted that the majority of errors of this type occur because stduents are close to the cutting score. Thus, the problem does not seem to be one that needs to be taken too seriously.

Receiving credit for non-mastered objectives on the posttest to the extent that the objectives are prerequisites to others in future units will interfere with the rate of learning at that point. This error seems to be less serious in terms of program efficiency if the objectives are terminal. Failing to receive credit for mastered objectives would seem to be less serious since the student could move through the remedial materials quickly and retake the test.

Since any decisions on the basis of curriculum-embedded test score results affect the student for only a limited amount of time and there exist checks on any decisions with the unit (or module) posttest, there is little concern for developing more appropriate testing decision guidelines at this level.

Assignment of Instructional Modes

An integral component of nearly every individualized instruction program is the feature whereby there exists several alternate instructional modes for the various units of instruction that can be assigned in some optimal way to students. In theory anyway, with IPI and PLAN, past performance and background aptitude variables are used to assist the students in

selecting the "best" mode of instruction. With Mastery Learning, this feature can be operationalized following the group-based instruction and the unit posttests. It is at this point that decisions on the proper corrective feedback techniques to use need to be made.

Investigators of the possible interactions between instructional methods and aptitudes are conducting what has been termed Aptitude-Treatment Interaction (ATI) research (Cronbach, 1967). Disappointing is the fact that while nearly all developers of individualized programs include this feature of utilizing ATI results in assigning instruction, there are few real demonstrations of significant interactions between aptitudes and instructional modes (Bracht, 1970; Cronbach and Snow, 1969). Authors such as Glaser (1972) have attempted to explain these results and suggest some new directions for this line of inquiry. However, it would appear that we are far from a "theory of instruction" to guide the instructional decision maker in the assignment of "optimal" instructional modes to students.

The benefits (assuming equal treatment costs) of the ATI classification scheme for improving the quality of instruction depend directly on the differences among the slopes of the regression lines for predicting criterion scores with different aptitude variables in the different instructional modes. The bigger the difference in slopes the greater is the potential benefit to the student for assigning one instructional mode or another. However, in looking at the overall benefits and losses of such a system, it would seem that the appropriate baseline for comparative purposes would need to be data derived from a traditional instructional program.

Final Year-End Assessment

This particular feature seems to be handled in much the same way in IPI and PLAN. Information is reported on the number and nature of units that a student has mastered. Little or no information is provided by the school to students and parents that could be used for norm-referenced assessment. In the mastery learning model, a score is reported to measure achievement on the year-long activities. Both norm-referenced and criterion-referenced interpretations are possible.

VI. Some Directions for Further Research

6.1 Concluding Remarks

A review of IPI, PLAN, and Mastery Learning programs as well as many other objective-based curriculum programs not reported in this paper reveals that there are many important questions remaining to be answered in regard to individual assessment models. In this concluding section a few of the more important problem areas are discussed.

In order to develop an instructional model that is sensitive to individual needs, abilities, interests, and goals in a way that will allow the student to maximize his learning, we need a theory of instruction. A theory of instruction should set down rules on the most efficient way of achieving knowledge (Bruner, 1964). This theory would provide guidelines on how to prescribe instruction to increase learning. One paper that addresses the problem is Groen and Atkinson (1966). Current reports on the related topic of aptitude-treatment interactions are by Cronbach and Gleser (1965), Cronbach and Snow, (1969), Bracht (1970), and Glaser (1972).

In making decisions on the basis of criterion-referenced test scores, one assumes a good match between items and the behavioral objectives they are intended to measure. To the extent that test items do not accurately measure the objectives, any decisions based on test performance will be inaccurate. To date a satisfactory methodology for item validation does not exist although several useful papers provide partial solutions (Dahl, 1971; Rovinelli and Hambleton, 1973).

A theory of criterion-referenced tests and measurements is also needed to guide the users of the tests in the context of programs

described here. This theory should probably be based on a threshold loss function rather than a squared-error loss function as has been done in classical test theory (Lord and Novick, 1968; Hambleton and Novick, 1973). This theory would include reliability, validity, test scoring, and item validation procedures for criterion-referenced tests. It would also provide guidelines and techniques for setting test length and cutting scores, and allocating testing time. A recent paper by Millman (1973) provides some excellent guidelines on this latter set of problems.

Another problem which has to be reckoned with for criterion-referenced tests is an instance of the bandwidth-fidelity issue (Cronbach and Gleser, 1965). When the total testing time is fixed and there is interest in measuring many competencies, one may be faced with the problem of whether to obtain very precise information about a small number of skills or less precise information about many more skills. Time allocation algorithms (analytical procedures for deciding how many items on a test should measure each objective) of a rather different kind than those presented by Woodbury and Novick (1968), and Jackson and Novick (1970) will be required. The problem of how to determine the number of items to measure each skill so as to maximize the percentage of correct decisions or some similar measure of overall decision-making accuracy on the basis of test results has yet to be resolved.

Estimation of mastery is a problem that is encountered frequently in individualized instructional programs. Bayesian methods have been suggested (Hambleton and Novick, 1973), but there has been no empirical demonstrations of their usefulness in this context nor are guidelines for the use of Bayesian methods available at the present time. Prior information for a Bayesian solution might include student mastery scores

on other skills covered on the test or student performance on skills measured on previous tests. (In the case of posttesting, pretest information could be used as the prior.) Also, just as data from other examinees can improve the precision of estimation of achievement in a norm-referenced testing situation for an individual (Lord and Novick, 1968), so perhaps the same can be done with criterion-referenced measurement problems.

Within many objective-based programs the strategy of branched testing would seem to be an appropriate technique, at least in situations where the objectives in a content area can be arranged into hierarchical sequences. Some of the practical problems have been resolved in the Pittsburgh IPI Program so that the technique can now be used on a limited basis. Nevertheless, many problems remain before adoption should or can proceed with other programs. For example, it would be necessary to develop a non-automated modified version of branched testing for schools without computers. Also, we need to know much more about starting places, step sizes, stopping rules, etc., before we can effectively use branched testing in an instructional setting.

References

Airasian, P. W. The role of evaluation in mastery learning. In J. H. Block, (Ed.), Mastery Learning: Theory and Practice. New York: Holt, Rinehart, and Winston, 1971.

Atkinson, R. C. Computer-based instruction in initial reading. In Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service, 1968.

Atkinson, R. C., & Wilson, H. A. (Eds.) Computer-assisted Instruction: A Book of Readings. New York: Academic Press, 1969.

Baker, F. B. Computer-based instructional management systems: A first look. Review of Educational Research, 1971, 41, 51-70.

Block, J. H. The effects of various levels of performance on selected cognitive, affective, and time variables. Unpublished doctoral dissertation, University of Chicago, 1970.

Block, J. H. (Ed.) Mastery Learning: Theory and Practice. New York: Holt, Rinehart, and Winston, 1971.

Block, J. H. Student evaluation: Toward the setting of mastery performance standards. Paper presented at the annual meeting of AERA, Chicago, 1972.

Bloom, B. S. (Ed.) Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York: D. McKay, 1956.

Bloom, B. S. Learning for mastery. Evaluation Comment, 1968, 1(2).

Bloom, B. S. Mastery learning. In J. H. Block (Ed.), Mastery Learning: Theory and Practice. New York: Holt, Rinehart, and Winston, 1971.

Bock, R. D., & Wood, R. Test theory. Annual Review of Psychology, 1972, 20, 193-224.

Bracht, G. H. Experimental factors related to aptitude-treatment interaction. Review of Educational Research, 1970, 40, 627-646.

Bruner, J. S. Some theorems on instruction stated with reference to mathematics. In E. R. Hilgard (Ed.), Theories of Learning and Theories of Instruction. (63rd National Society for the Study of Education Yearbook, Part I), Chicago: University of Chicago Press, 1964, 306-335.

Carroll, J. B. A model of school learning. Teachers College Record, 1963, 64, 723-733.

Carroll, J. B. Problems of measurement related to the concept of learning for mastery. Educational Horizons, 1970, 48, 71-80.

Cooley, W. W., & Glaser, R. The computer and individualized instruction. Science, 1969, 166, 574-582.

Cox, R. C., & Boston, M. E. Diagnosis of Pupil Achievement in the Individually Prescribed Instruction Project. (Working Paper 15.) Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1967.

Cronbach, L. J. How can instruction be adapted to individual differences? In R. M. Gagné (Ed.), Learning and Individual Differences. Columbus, Ohio: Charles E. Merrill, 1967.

Cronbach, L. J., & Gleser, G. C. Psychological Tests and Personnel Decisions. (2nd ed.) Urbana, Ill.: University of Illinois Press, 1965.

Cronbach, L. J., & Snow, R. E. Individual Differences in Learning Ability as a Function of Instructional Variables. (USOE Final Report, Contract No. OEC 4-6-061269-1217.) Stanford, California: Stanford University, School of Education, 1969.

Dahl, T. Toward an evaluative methodology for criterion-referenced measures: Objective-item congruence. Paper presented at the annual meeting of the California Educational Research Association, San Diego, 1971.

De Vault, M. V., Kriewall, T. E., Buchanan, A. E., & Quilling, M. R. Teacher's manual: Computer Management for Individualized Instruction in Mathematics and Reading. Madison, Wisconsin: University of Wisconsin, Research and Development Center for Cognitive Learning, 1969.

Dunn, J. A. The PLAN approach to curriculum definition. Education, 1970, 90, 221-226.

Ebel, R. L. Criterion-referenced measurements: Limitations. School Review, 1971, 69, 282-288.

Ferguson, R. L. The development, implementation, and evaluation of a computer-assisted branched test for a program of individually prescribed instruction. Unpublished doctoral dissertation, University of Pittsburgh, 1969.

Ferguson, R. L. Computer Assistance for Individualizing Measurement. Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1971.

Flanagan, J. C. Functional education for the seventies. Phi Delta Kappan, 1967, 49, 27-32.

Flanagan, J. C. Program for learning in accordance with needs. Psychology in the Schools, 1969, 6, 133-136.

Flanagan, J. C. Individualizing education. Education, 1970, 90, 191-206.

Flanagan, J. C., Dailey, J. T., Davis, F. B., Goldberg, I., Heyman, C. A., Jr., Orr, D. B., & Shaycraft, M. F. The American High School Student. (USOE Final Report, Cooperative Research Project No. 635.) Pittsburgh: University of Pittsburgh, 1964.

Gagné, R. M. The Conditions of Learning. New York: Holt, Rinehart, and Winston, 1965.

Gibbons, M. What is individualized instruction? Interchange, 1970, 1, 28-52.

Glaser, R. Instructional technology and the measurement of learning outcomes. American Psychologist, 1963, 18, 519-521.

Glaser, R. Adapting the elementary school curriculum to individual performance. In Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service, 1968.

Glaser, R. Evaluation of instruction and changing educational models. In Wittrock, M. C. and Wiley, D. E. (Eds.), The Evaluation of Instruction. New York: Holt, Rinehart, and Winston, 1970.

Glaser, R. Individuals and learning: The new aptitudes. Educational Researcher, 1972, 6, 5-13.

Glaser, R., & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.), Educational Measurement. Washington: American Council on Education, 1971, 625-670.

Groen, G. J., & Atkinson, R. C. Models for optimizing the learning process. Psychological Bulletin, 1966, 66, 209-320.

Hambleton, R. K., & Gorth, W. P. Criterion-referenced testing: Issues and applications. Center for Educational Research Technical Report No. 13. Amherst, Mass.: University of Massachusetts, Amherst, 1971.

Hambleton, R. K., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10, in press.

Heathers, G. Overview of innovations in organization for learning. Interchange, 1972, 3, 47-68.

Hieronymous, A. N. Today's testing: What do we know how to do? In Proceedings of the 1971 Invitational Conference on Testing Problems. Princeton, N. J.: Educational Testing Service, 1972.

Hsu, T. C., & Carlson, M. Oakleaf School Project: Computer-assisted Achievement Testing. (A Research Proposal.) Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1972.

Jackson, P. H., & Novick, M. R. Maximizing the validity of a unit-weight composite as a function of relative component lengths with a fixed total testing time. Psychometrika, 1970, 35, 333-347.

Jung, S. M. Guidance testing and the individualized program of studies. Education, 1970, 90, 224-231.

Kim, H. The mastery learning project in the middle schools. In J. H. Block (Ed.), Mastery Learning: Theory and Practice. New York: Holt, Rinehart, and Winston, 1971. (Abstract)

Lindvall, C. M., & Cox, R. The role of evaluation in programs for individualized instruction. In National Society for the Study of Education, Educational Evaluation: New Roles, New Means, Sixth-Eighth Yearbook, Part II, 1969, 156-188.

Lindvall, C. M., Cox, R. C., & Bolvin, J. O. Evaluation as a tool in curriculum development: The IPI evaluation program. AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally, 1970.

Livingston, S. A. Criterion-referenced applications of classical test theory. Journal of Educational Measurement, 1972, 9, 13-26.

Lord, T. M. Some test theory for tailored testing. In W. H. Holtzman (Ed.), Computer-assisted Instruction, Testing and Guidance. New York: Harper and Row, 1970.

Lord, F. M., & Novick, M. R. Statistical Theories of Mental Test Scores. Reading, Mass.: Addison-Wesley, 1968.

Mayo, S. T. Mastery learning and mastery testing. Measurement in Education, 1970, 1(3).

Millman, J. Passing scores and test length for domain-referenced measures. The Review of Educational Research, 1973, 43, 205-216.

Popham, W. P. Curriculum materials. Review of Educational Research, 1969, 39, 319-338.

Popham, W. J., & Husek, T. R. Implications of criterion-referenced measurement. Journal of Educational Measurement, 1969, 6, 1-9.

Rhetts, J. E. The impact of student learning style on curriculum assignment and performance in the PLAN program of individualized instruction. Education, 1970, 90, 248-251.

Rovinelli, R., & Hambleton, R. K. Some Procedures for the Validation of Criterion-Referenced Test Items. (Final Report) Albany, New York: Bureau of School and Cultural Research, New York State Education Department, 1973.

Scriven, M. The methodology of evaluation. In R. Stake (Ed.), Perspectives of Curriculum Evaluation. Chicago: Rand McNally, 1967.

Smith, H. A. Curriculum development and instructional materials. Review of Educational Research, 1969, 39, 397-414.

Suppes, P. The uses of computers in education. Scientific American, 1966, 215, 206-221.

Washburne, C. W. Educational measurements as a key to individualizing instruction and promotions. Journal of Educational Research, 1922, 5, 195-206.

Weisgerber, R. A. (Ed.) Developmental Efforts in Individualizing Learning. New York: Peacock, 1971.

Weisgerber, R. A., & Rahmlow, H. F. Strategies for learning in PLAN. In R. A. Weisgerber, (Ed.), Developmental Efforts in Individualized Learning. Itasca, Ill.: Peacock, 1971.

Wilhelms, F. T. The curriculum and individual differences. In National Society for the Study of Education, Individualizing Instruction, Sixty-first Yearbook, Part I, 1962, 72-82.

Wittrock, M. C., & Wiley, D. E. (Eds.) The Evaluation of Instruction. New York: Holt, Rinehart, and Winston, 1970.

Woodbury, M. A., & Novick, M. R. Maximizing the validity of a test battery as a function of relative test lengths for a fixed total testing time. Journal of Mathematical Psychology, 1968, 5, 242-259.

Appendix Number 2

to

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

Appendix Number 2

to

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

ACT TECHNICAL BULLETIN NO. 17

A Primer on Decision Analysis for Individually Prescribed Instruction

by

Charles E. Davis
The University of Iowa

James Hickman
The University of Wisconsin

and

Melvin R. Novick
The American College Testing Program
and
The University of Iowa

The Research and Development Division
The American College Testing Program
P. O. Box 168, Iowa City, Iowa 52240
October, 1973

A Primer on Decision Analysis for Individually Prescribed Instruction[1]

Charles E. Davis
The University of Iowa

James Hickman
The University of Wisconsin

and

Melvin R. Novick

The American College Testing Program
and
The University of Iowa

1. Introduction

If you are lucky, and certain that your luck will hold, you should read no further. Our subject is decision making and those who are inherently lucky will have no need to attempt to take advantage of a logical system for decision making. However, we feel that logical thought has been successful in so many areas of human activity that it deserves a chance even in those areas where arbitrary rules or intuition seem to have prevailed, and perhaps been moderately successful. In particular, we feel that the implementation of Individually Prescribed Instruction can be given a greater payoff if a coherent system of decision making is incorporated into the instructional sequence to provide a supplement to the experience-honed judgment of the classroom teacher.

In the currently popular language of systems engineering, the decision making process might be viewed as a black box. The black box contains an input hole for prior information about the environment in which the decision will be made and evaluated, a second hole for new experimental results designed

[1]The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Dep:r'ment of Health, Education, and Welfare, Melvin R. Novick, Principal Investigauor. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy. We are grateful to Nancy S. Peterson for carefully reading and correcting this manuscript.

explicitly to cast more light on the uncertain environment, and finally a hole for the decision maker's preferences for the possible consequences of decisions. The output of this black box is a decision or action that will ultimately lead to a distribution of consequences that will, hopefully, be high on the decision maker's preference scale. Our business will be to get enough of a look at what goes on inside the black box so that we can see how the teacher and student can use prior information, experimental results, and preferences in a reasonable way to generate decisions having favorable consequences.

Building an all-purpose black box for decision making has been a major project in recent years. Contributors to the project have included economists, mathematicians, philosophers, psychologists, and statisticians. It would take a major treatise to adequately trace this project. We must, therefore, limit our survey to the major references in psychology and education relevant to our task. Those elements of decision analysis which require that existing information about the process under study be quantified in order that new experimental information may be coherently combined with the existing information, have already been persuasively presented in the literature of psychology and education [Edwards, Lindman, and Savage (1963)]. However, the requirement that preferences for various possible distributions of consequences be formulated coherently and expressed as a numerical valued utility function, has not been emphasized in the literature of education with the singular exception of the prophetic text, Psychological Tests and Personnel Decisions, by Cronbach and Gleser (1957, 1965).

Many decision makers in education may feel that business managers, with the market at hand to evaluate the outcomes associated with their decisions, are in a more favorable position to make value comparisons among the

distributions of consequences than they are. For example, the values attached to the consequences of a decision to implement a specific education program, or to advance or retain a particular student at a given level, seem to them to defy simple quantification. Such decisions appear to have so many possible ramifications that to formulate preferences with any degree of consistency is simply impossible.

Nevertheless, it is our position that most data-collecting activities in education are designed to influence decisions. In fact, to simply collect data without the objective of ultimately modifying a course of action, would seem wasteful. We further assert that despite all the perplexities, decisions are regularly made in education based on an informal mixture of recently collected information, prior information, and the preferences of the decision maker. If these decisions are to be rational in the sense that they are derived from a logical program for decision making that provides for the input of prior information, new data, and preferences, we claim that decision theory is required. There is no magic in the formal structure of decision theory. The theory contributes only what mathematics does to any problem; an orderly, systematic, and precise framework for formulating a problem, plus the economy of mathematical reasoning in tracing the consequences of the formulation.

The main difficulty in implementing decision theory arises from the necessity to quantify basically subjective or personalistic quantities, and this difficulty is real. However, we believe that within the framework of some of the newer and very highly structured modes of instruction, it is possible to provide relatively simple yet conclusively meaningful methods of decision analysis. The decision machinery that we shall build will be appropriate for a wide band of decision problems. However, we will consistently illustrate the ideas with examples from individualized instructional procedures.

2. Fundamental Ideas and Notation

The main purpose of this section is to illustrate the rudiments of decision analysis with a variety of straightforward examples. Thus, we begin Section 2.1 with a rather trivial example. In later sections, this example will be modified and extended to make it both more realistic and interesting.

2.1 The Basic Example

The environment of our example is very simple. A student has either mastered or not mastered the topics in the current unit of his individually prescribed instructional program. The state of the student being a nonmaster will be denoted by θ_1 and the state of his being a master will be denoted by θ_2. Two actions are open to the decision maker, who could be the teacher or even, in some arrangements, the student himself. The decision may be to retain the student for additional work at the present level of his prescribed program. This "retain" decision will be denoted by d_1. The complement of this decision is to advance the student to the next unit. This decision will be denoted by d_2.

We now identify the three basic inputs into the decision analysis.

(a) Prior information. In Individually Prescribed Instruction, a student begins a unit of instruction only when he is deemed to be prepared for that unit. For this reason, the variation in posttest results tend to be relatively small, except to the extent that they are due to sampling variation. Therefore, the main input of prior information will involve our beliefs about the relative success of the instructional unit with qualified entrants. Thus, on the basis of an examination of the success which other students have had on this training unit, and before administering a test at the end of the present unit, the decision maker assigns prior probabilities to the two possible states.

Table 1

State of the Student

		θ_1 (nonmaster)	θ_2 (master)
Prior Probability	$\rho(\theta)$	.4	.6

This information is relevant to the decision under consideration and to ignore it is to waste useful facts.

The symbol $\rho(\theta)$ denotes the prior probability mass function. This symbol will continue to play the same role throughout our discussion.

(b) Experimental plan. The decision maker may ask the student one test question to learn more about which state prevails. The result of this short test will be denoted by X . If the student answers correctly, X = 1, and if he answers incorrectly, X = 0 . The probability assignments to these two outcomes, if the mastery state is known, is given in the following short table where $p(x|\theta) = Pr[X = x|\theta]$ is the probability mass function of correct (x = 1) and incorrect (x = 0) responses given that the mastery evel is known to be θ .

Table 2

$p(x|\theta)$

	x = 0 (wrong)	x = 1 (right)
θ_1 (nonmaster)	.8	.2
θ_2 (master)	.2	.8

This table states that if a person is a master, then the probability that he will give a correct answer is .8 and the probability that he will give a wrong response is .2. If he is a nonmaster, the probabilities are reversed. Note that those are probabilities of experimental outcomes given the true state of the student (master, nonmaster).

(c) Preferences. The decision maker can make one of the two types of errors. If he retains the student at the current level when, in fact, the student is a master, the student will probably repeat the current unit with only minimal gain. On the other hand, if the student is advanced when he has not mastered the topics on the current level, ultimately he may have to repeat both the current level and the one to which he had prematurely been advanced. With these facts in mind, the decision maker designates the nonnegative loss function $L(d, \theta)$ defined in Table 3.

Table 3

$L(d, \theta)$

	θ_1 (nonmaster)	θ_2 (master)
d_1 (retain)	0	1
d_2 (advance)	2	0

In specifying a loss function of this type, the decision maker assumes that no loss occurs if a correct decision is made and that the loss associated with advancing a nonmaster is twice that associated with retaining a master. We do not suggest that this simple loss function is appropriate in all or even any situations. However, the simple assumptions seem realistic enough to maintain our attention for a while.

Three reasonable decision rules for selecting decisions d_1 or d_2, after observing the test score, are possible.

Table 4

Rule	$x = 0$ (wrong)	$x = 1$ (right)
$\delta_1(x)$	d_1	d_1
$\delta_2(x)$	d_1	d_2
$\delta_3(x)$	d_2	d_2

In this trivial example, decision rules $\delta_1(x)$ and $\delta_3(x)$ really tell us to ignore prior information and the current test result; in the case of $\delta_1(x)$ to take action d_1 (retain) invariably, and in the case of $\delta_3(x)$ to stick with action d_2 (advance). Nevertheless, each is a serious candidate because certain loss structures and prior probability distributions could make one of them the preferred decision rule.

There are two ways that the analyses carried on within the black box labeled "decision process" have historically been organized. The first, which is called normal form analysis, involves a three step process. (1) Compute the average or expected loss for each ordered pair (δ_i, θ_j) composed of a decision rule δ_i and a state parameter value θ_j. This averaging is performed with respect to the probability distribution of possible experimental outcomes, $p(x|\theta_j)$.

The expected value that emerges from this computation is called a risk function and it is denoted by $R(\delta_i, \theta_j)$. As the notation suggests, it is a function of both the decision rule δ_i and the state parameter θ_j. In symbols, we have

$$R(\delta_i, \theta_j) = \sum_x L[\delta_i(x), \theta_j]p(x|\theta_j)$$

where i indexes the various decision rules and j identifies the various values of the state parameter θ. Once we have the risk function, we have the expected loss for each possible (δ_i, θ_j) pair, and can possibly decide that a decision rule is good (bad) if its risk is small (large) for all values of θ_j. (2) This kind of analysis, however, is typically inconclusive so we must compute the average or expected value of the risk function for each decision rule. This averaging is done with respect to the prior probability distribution of the state parameter $\rho(\theta)$. This

expected value is called the Bayes risk function and will be denoted by $r(\delta_i, \rho)$, where the symbol ρ is inserted to stress that the Bayes risk depends on the prior probability distribution. This recipe may be expressed in symbols as

$$r(\delta_i, \rho) = \sum_{\theta} R(\delta_i, \theta_j)\rho(\theta_j) \quad .$$

The effect of this averaging is to weight the effect of a decision rule highly for those values of θ_j that we think, apriori, are highly probable and, hence, important to consider. (3) With this computation completed for each decision rule, we select and subsequently use the decision rule that has the smallest Bayes risk.

Now let us carry out this three step normal form analysis using the losses and prior probabilities from our example. First, we will compute the risk of decision rule δ_2 when the true state of the person is θ_1. Recall that with δ_2, we retain the student if $x = 0$ and advance him if $x = 1$. The risk is

$$R(\delta_2, \theta_1) = \sum_{x} L[\delta_2(x), \theta_1]p(x|\theta_1)$$

$$= L[\delta_2(0), \theta_1]p(0|\theta_1) + L[\delta_2(1), \theta_1]p(1|\theta_1) \ ,$$

i.e., the risk of using decision rule δ_2 when the true state is θ_1 is the simple average or expected loss for the (δ_2, θ_1) pair. The expectation is performed with respect to the probability distribution $p(x|\theta_1)$ of the two possible test scores $x = 0$ and $x = 1$, when it is given that the student is a nonmaster (i.e., $\theta = \theta_1$). Let us evaluate this risk. From Table 4, we know that $\delta_2(0) = d_1$ and $\delta_2(1) = d_2$. Thus,

$$R(\delta_2, \theta_1) = L(d_1, \theta_1)p(0|\theta_1) + L(d_2, \theta_1)p(1|\theta_1) \quad .$$

From Table 3, we have $L(d_1, \theta_1) = 0$ and $L(d_2, \theta_1) = 2$; from Table 2, we have $p(0|\theta_1) = .8$ and $p(1|\theta_1) = .2$. Thus, using the symbol * to indicate multiplication, we have

$$R(\delta_2, \theta_1) = 0 * .8 + 2 * .2 = .4 ,$$

i.e., when the true situation is θ_1, decision rule δ_2 has a risk (expected or average loss) of .4. Similar computations have been made for the remaining (δ_i, θ_j) pairs. The computations have been summarized in Table 5.

Table 5

$$R(\delta_i, \theta_j) = \sum_x L[\delta_i(x), \theta_j]p(x|\theta_j)$$

State	Result	$p(x\|\theta)$	$\delta_1(x)$	$\delta_2(x)$	$\delta_3(x)$
θ_1	x = 0	.8	0 * .8	0 * .8	2 * .8
	x = 1	.2	0 * .2	2 * .2	2 * .2
		$R(\delta_i, \theta_1)$	0	.4	2.0
θ_2	x = 0	.2	1 * .2	1 * .2	0 * .2
	x = 1	.8	1 * .8	0 * .8	0 * .8
		$R(\delta_i, \theta_2)$	1.0	.2	0

If we *knew* that θ_1 (nonmaster) prevailed, decision rule $\delta_1(x)$ which always retains the student would minimize the risk function. It has an expected loss of zero which is as good as one can do. If we *knew* that θ_2 (master) prevailed, decision rule $\delta_3(x)$ which always advances the student would minimize the risk function. Again this rule would have no risk. 143

Of course, if we knew which state prevailed there would be no need to apply

decision analysis to the problem. At this point, none of the three decision rules are judged to be uniformly superior, i.e., superior for every state of nature. If one were, we would certainly adopt it, but none is, so we must find some way of choosing the best decision rule. We, therefore, move to Step (2) in normal form analysis before identifying the winner. The computation of the Bayes risk must now be made for each of the three decision rules. The idea is simply this. We do not know the true state of the person, but we do have a prior opinion concerning the true state. Therefore, it makes sense to average the risk for each decision rule with respect to our prior opinions, in effect, to put more weight on those values of θ that seem more probable to us. For example, to compute the Bayes risk $r(\delta, \rho)$ for δ_2 (the average risk with respect to our prior probabilities for θ_1 and θ_2), we compute the following:

$$r(\delta_2, \rho) = R(\delta_2, \theta_1)\rho(\theta_1) + R(\delta_2, \theta_2)\rho(\theta_2) \quad .$$

Substituting values from Table 5 and Table 1, we have

$$(.4 * .4) + (.2 * .6) = .28 \ .$$

The computations of the Bayes risk for δ_1 and δ_3 are also easily made and are given, together with those for δ_2, in Table 6.

Table 6

θ	$\rho(\theta)$	$R(\delta_1, \theta)\rho(\theta)$	$R(\delta_2, \theta)\rho(\theta)$	$R(\delta_3, \theta)\rho(\theta)$
θ_1	.4	0 * .4	.4 * .4	2 * .4
θ_2	.6	1 * .6	.2 * .6	0 * .6
	$r(\delta_1, \rho)$	.6	.28	.80

The final step in normal form analysis directs us to select action rule δ_2 because it has the smallest Bayes risk. Thus, we have stated and exhibited in detail a precise and coherent procedure for decision-making in the presence of uncertainty. We have further demonstrated a simple application in the context of Individually Prescribed Instruction. To see that the choice of decision rule really depends on prior probabilities, the reader should redo step (3) in the analysis first with $\rho(\theta_1) = .1$, $\rho(\theta_2) = .9$, and then with $\rho(\theta_1) = .9$, $\rho(\theta_2) = .1$.

A second way of organizing the analysis within the decision process black box is called <u>extensive form analysis</u>. Since this type of analysis has some computational advantages over the normal form, our subsequent illustrations will, with one exception, employ extensive form analysis.

Extensive form analysis also involves a three step process. (1) Determine the posterior probability distribution of the state parameter θ, given the experimental result x. That is, we must determine $p(\theta|x) = p(x|\theta)\rho(\theta)/p(x)$, where $p(x) = \sum_{\theta} p(x|\theta)\rho(\theta)$ is the unconditional probability mass function of X. The posterior distribution $p(\theta|x)$ summarizes our knowledge and beliefs about θ, incorporating both our prior beliefs and the sample information. (2) Compute the expected value of the loss function for each decision rule with respect to the posterior distribution of θ. That is, we must compute $\sum_{\theta} L(\delta_i, \theta)p(\theta|x)$ for each decision rule. (3) Select the decision that will yield the smallest posterior expected loss in Step (2).

The advantage of extensive form analysis arises in Step (2) and is a bit hard to appreciate when expressed only in words. The heart of the matter is that Step (2) does <u>not</u> have to be carried out for every possible value of X. If we adopt this system, we can wait and perform Step (2) only

for the result x that is actually observed. Once x is observed, the decision rule δ_i specifies the decision d_i to be made, and the losses under each state θ can be taken immediately from Table 3.

The amazing thing is that under very general conditions, normal and extensive form analysie will lead to the same decision. This point will be illustrated in our example. Later, a mathematical argument will be presented for those who can only be persuaded by such demonstrations (Raiffa and Schlaifer, 1961, p. 15).

The first step in extensive form analysis requires us to determine the form of the posterior probability distribution of θ. For our example, we will do more and actually exhibit the two possible posterior distributions. But first, we need to exhibit the joint probability distribution of X and θ, and the marginal distribution of X. These probabilities are given in Table 7 where the entries in the body of the table are the joint probabilities, and the entries in the margins are the marginal probabilities. In computing Table 7, we have used $\rho(\theta)$ from Table 1 and $p(x|\theta)$ from Table 2.

Table 7

$p(x|\theta)\rho(\theta) = p(x, \theta)$

	x = 0	x = 1	$\rho(\theta)$
θ_1	.32	.08	.40
θ_2	.12	.48	.60
p(x)	.44	.56	

Then the posterior probability distribution of θ for given values of x is given by Bayes Theorem $p(\theta|x) = p(x, \theta)/p(x)$. These conditional probabilities for θ given x are summarized in Table 8. Note that the conditional distribution of θ given x = 0 is very different from that given x = 1.

Table 8

$p(\theta|x)$

	$x = 0$	$x = 1$
θ_1	.73	.14
θ_2	.27	.86
	1.00	1.00

The second step in extensive form analysis calls for the computation of the average or expected value of the loss function for each decision rule with respect to the posterior distribution of θ [i.e., we must compute $\sum_j L(\delta_1, \theta_j)p(\theta_j|x)$]. We will carry out this computation for each of the three decision rules specified in Table 4. Suppose $x = 0$. Then from Table 8, we see that $p(\theta_1|x = 0) = .73$ and $p(\theta_2|x = 0) = .27$. If we adopt δ_1, then we shall make decision d_1 when $x = 0$ (see Table 4). Thus, if $\theta = \theta_1$, our loss will be zero, and if $\theta = \theta_2$, our loss will be one (see Table 3). Therefore, our average or expected loss given $x = 0$ is

$$\begin{aligned}\mathcal{E}\{L[\delta_1(0), \theta]|x = 0\} &= L[\delta_1(0), \theta_1]p(\theta_1|x = 0) \\ &\quad + L[\delta_1(0), \theta_2]p(\theta_2|x = 0) \\ &= (0 * .73) + (1 * .27) = .27\end{aligned}$$

as given in the first column of Table 9.

Table 9

$\mathcal{E}\{L[\delta_i(x), \theta]|x\}$

	x = 0			x = 1		
	δ_1	δ_2	δ_3	δ_1	δ_2	δ_3
θ_1	0 * .73	0 * .73	2 * .73	0 * .14	2 * .14	2 * .14
θ_2	1 * .27	1 * .27	0 * .27	1 * .86	0 * .86	0 * .86
	.27	.27	1.46	.86	.28	.28

In this table, similar computations are made for each of the three decision rules for both x = 0 and x = 1. Observe that decision rule $\delta_2(x)$ produces the smallest expected loss for each value of x and may be judged as the optimum decision rule. However, if x = 0, rule δ_1 is equally as good, and if x = 1, rule δ_3 is equally as good. More importantly, however, this observation indicates that the second and third steps can be simplified still further. Note that if x = 0, $\delta_1(x)$ and $\delta_2(x)$ lead to the same decision (d_1) and, consequently, must necessarily have the same expected loss. What this emphasizes is that once we know x, we are really interested only in the best <u>decision</u> (d_1 or d_2), rather than the decision rule (δ_1, δ_2, δ_3) that will lead to the best decision. Consequently, we need only compute the expected or average value of the loss function for the available decisions. That <u>decision</u> (not decision rule) with the smallest posterior loss will then be selected. This is done in Table 10.

Table 10

$$\sum_{\theta} L(d, \theta) p(\theta | x)$$

	x = 0		x = 1	
	d_1	d_2	d_1	d_2
θ_1	0 * .73	2 * .73	0 * .14	2 * .14
θ_2	1 * .27	0 * .27	1 * .86	0 * .86
	.27	1.46	.86	.28

Our action rule is the same as indicated previously; if we observe $x = 0$, we will take action d_1 (retain), and if we observe $x = 1$, we will take action d_2 (advance). This is, in effect, the same as adopting decision rule δ_2 . As pointed out previously, in extensive form analysis, only that half of Table 10 which corresponds to the actual result observed ($x = 0$ or $x = 1$) needs to be calculated.

So far, we have illustrated two forms of analysis which we claim will lead to identical decisions. These forms combine prior information, experimental results, and preferences using the machinery of probability theory to trace the consequences of the inputs. The example that helped to illustrate these ideas was kept at a trivial level so that the arithmetic would not obscure the essence of the process.

2.2 The Formal Structure of Decision Theory

Although mathematical symbols may seem forbidding, they are an indispensable tool in conveying ideas precisely. Therefore, both for completeness and for precision we will retrace the key ideas of decision analysis relying on symbols rather than numbers. Although integral signs, $\int$, will be used, we remind our readers that this is merely a continuous analogue of the summation operator, Σ .

We let $L[\delta(x), \theta)]$ be a nonnegative loss function defined on the set of pairs of the parameter values, δ and θ. The symbol $\delta(x)$ is used to emphasize the fact that the decision rules usually depend on observed values of the random variable X; the observation being obtained to elicit information about θ. In some contexts it may help to think of θ as the mean of some probability distribution, and $\delta(x)$ as an estimate of θ. Then $L[\delta(x), \theta]$ is the nonnegative numerical value of the loss associated with using $\delta(x)$ as an estimate of the mean, when, in fact, the mean is θ.

In the first stage of normal form analysis, we need the expected or average value of this loss function, where expectation is taken with respect to the conditional distribution of X given θ. This expected value was referred to as the risk function and was denoted $R(\delta, \theta)$. That is

$$R(\delta, \theta) = \int L[\delta(x), \theta]p(x|\theta)dx .$$

For this integral, the symbol $p(x|\theta)$ denotes the probability density of the random variable X, given the value of the state parameter θ. If X is a discrete random variable, $p(x|\theta)$ is to be interpreted as a probability mass function, and integration is then replaced by summation. As the symbols suggest, the risk function is similar to the loss function. The major difference rests on the fact that the risk function no longer depends upon the observed value of our experimental variable X; X has been averaged out. This fact is emphasized by the absence of an x in our notation for the risk function, $R(\delta, \theta)$.

If one adopts a Bayesian view of statistics, and one is compelled to embrace this view if he accepts any of several comprehensive axiom systems for decision making, it becomes necessary to quantify previous

or collateral information about the state parameter θ, which we summarize in $\rho(\theta)$. In the second stage of normal form analysis, the expected or average value of the risk function from stage one is then calculated with respect to this prior distribution, $\rho(\theta)$. We referred to this expected risk as the Bayes risk associated with δ, and denoted it, $r(\delta, \rho)$. And so we have,

$$r(\delta, \rho) = \int R(\delta, \theta)\rho(\theta)d\theta$$
$$= \int\{\int L[\delta(x), \theta]p(x|\theta)dx\}\rho(\theta)d\theta \quad .$$

Then, the third and final stage of the normal form analysis consisted of selecting the decision rule which minimizes the Bayes risk.

As was pointed out above, extensive form analysis follows a slightly different route, but under rather general conditions, leads one to the same decision. In extensive form analysis, we begin by evaluating the posterior distribution of θ

$$p(\theta|x) = \frac{\rho(\theta)p(x|\theta)}{p(x)} \quad ,$$

and then determine the expected loss with respect to this posterior distribution. In the continuous case, this expected or average loss can be represented by the integral,

$$\mathcal{E}_{\theta|x}\{L[\delta(x), \theta]\} = \int L[\delta(x), \theta]p(\theta|x)d\theta \quad .$$

Naturally, in the discrete case, we merely replace the integration operation with that of summation. The decision making criterion in extensive form analysis is then to choose that decision rule $\delta(x)$, which minimizes this expected loss. The reader should take special note at this point, that

although this expected loss depends upon the random variable X, it needs to be evaluated only for that value x which is actually obtained.

In order to make clear the relationship between normal form and extensive form analyses, let us compare the decision criteria in the two cases. Under normal form analysis we are to choose the decision which minimizes the Bayes risk,

$$r(\delta, \rho) = \int\{\int L[\delta(x), \theta]p(x|\theta)dx\}\rho(\theta)d\theta \quad .$$

Now, if these integrals are suitably well behaved, we may interchange the order of integration and so write the Bayes risk in the form,

$$r(\delta, \rho) = \int\{\int L[\delta(x), \theta]p(x|\theta)\rho(\theta)d\theta\}dx \quad .$$

Since by Bayes theorem $p(x|\theta)\rho(\theta) = p(\theta|x)p(x)$, we may rewrite this last equation in the form

$$r(\delta, \rho) = \int\{\int L[\delta(x), \theta]p(\theta|x)d\theta\}p(x)dx \quad . \qquad (2.1)$$

The observant reader will of course already have recognized the integral in brackets as the expected loss which must be minimized under the extensive form approach. This equation illustrates the crucial difference between the two approaches. Extensive form analysis chooses the decision rule which minimizes the expected loss for the <u>particular value of x observed</u>. In contrast, normal form analysis chooses the decision rule which minimizes the average of those expected losses <u>for all possible values of x</u>. Clearly, if one particular decision rule minimizes the expected loss criterion of

extensive form analysis for every x, then the average of those expected losses under that decision rule must also be a minimum. In this case, it is clear that the decision taken under extensive form analysis will coincide precisely with that taken under normal form analysis. However, in those instances where extensive form analysis apparently leads to different decision rules depending upon the value of x observed, the equivalence of the two approaches may not be obvious.

Although the possible non-equivalence of these two approaches may seem to pose a dilemma for users of decision theory, in fact, it is a non-problem which has been set merely for the pedagogical purpose of underscoring some fundamental differences in the two approaches. Consider, if you will, the following. Since we are admitting for consideration all reasonable decision rules, we must allow that rule $\delta^*(x)$ which, for each x, minimizes the expected loss criterion of extensive form analysis. Since $\delta^*(x)$ will be selected by the extensive form approach irrespective of the value of x obtained, it will also be chosen under the normal form procedure. How do we construct $\delta^*(x)$? We do this in a straightforward operational manner: We use extensive form analysis for the x obtained and choose that <u>decision</u> (not decision rule) which minimizes the expected loss.

For a concrete example of the relationship between normal form and extensive form analysis, consider Table 11 which summarizes the expected losses under extensive form analysis for the three decision rules. This table is merely a modified version of Table 9.

Table 11

$$\mathcal{E}_{\theta|x}\{L[\delta_1(x),\ \theta]\}$$

	δ_1	δ_2	δ_3
x = 0	.27	.27	1.46
x = 1	.86	.28	.28

We found in Table 7 that the marginal distribution of X is given by pr(X = 0) = .44 and pr(X = 1) = .56. Thus, as indicated in Equation (2.1), the Bayes risk associated with decision δ_1 is given by the weighted average of the entries in the first column of Table 11, where the weights are the marginal probabilities of X . And so we have,

$$r(\delta_1,\ \rho) = (.27 * .44) + (.86 * .56) = .6$$

as we saw in Table 6.

2.3 Extensive Form Analysis with a Continuous Posterior

Let us now turn to an example of extensive form analysis which uses a continuous model density. This example is only a slight modification of that used previously; the primary difference being that we now assume that both the state parameter θ and the random variable X are continuous. It should be noted that this example is merely a reformulation of one considered by Hambleton and Novick (1972). As before, two decisions are open to the decision maker who is guiding a student through an ordered sequence of instructional units. At the end of each unit, the decision maker, based on his knowledge of the student's past performance, the performance of similar students, and current test results, must decide

to advance the student to the next unit in the sequence or to retain him at the present level. If the decision maker knows the student's mastery level (θ), he would be willing to advance the student if $\theta \geq \theta_o$ and to retain him if $\theta < \theta_o$. Thus, the number θ_o is a cutoff (or selection) point on the mastery scale with respect to the actions advance or retain.

In selecting θ_o, careful consideration of the objectives of the training program, and previous experience with the training and evaluation materials must prevail. If for example, θ_o is intended merely to give at least an even chance of completing the next lesson, θ_o might be set equal to that level of functioning which has historically had a 50% success rate on the next unit. If, on the other hand, the decision maker is very concerned about the ill-effects of the frustration of a poorly prepared student reading advanced material, perhaps θ_o ought to be somewhat higher. In any case, once θ_o is specified for the test, prior and collateral information about the student will be combined with the test result (x) for the purpose of estimating θ.

Assume that for this two-action (advance or retain) problem, the decision maker specifies a threshold (or step) loss function which can be described by the following table: (Compare with Table 3)

Table 12

$L(d_i, \theta)$

	$\theta < \theta_o$	$\theta \geq \theta_o$
d_1 (retain)	0	b
d_2 (advance)	c	0

where b and c are both nonnegative. In the literature, losses associated with falling into the lower left cell of this table are frequently referred to as arising from "false-positives", since the decision maker has wrongly presumed that the parameter θ lies in the region which has positive ethical value (i.e., has wrongly presumed that $\theta \geq \theta_o$). Similarly, losses associated with falling in the upper right cell are commonly referred to as arising from "false-negatives". The nonnegative numbers, b and c, reflect the cost of making these two types of errors. Because the decision envisaged in this example is rather local, affecting only one step in a program which is only a small part of the student's total learning experience, a massive effort to determine b and c exactly would seem inappropriate. In some cases, it might seem reasonable to assume that $c/b = 2$, if for example, a false positive results in repeating two steps in the sequence, as compared with the repetition of only one for a false negative.

Following the general scheme for extensive form analysis outlined earlier, our goal is to determine the action which will minimize the expected or average posterior loss

$$\int L(d_i, \theta)p(\theta|x)d\theta \quad .$$

This integral is equal to $b\{Pr[\theta \geq \theta_o|x]\}$ if $i = 1$ and is equal to $c\{Pr[\theta < \theta_o|x]\}$ if $i = 2$. Therefore, we may minimize this expected loss by making decision d_1 for those values of x such that

$$b\{Pr[\theta \geq \theta_o|x]\} < c\{Pr[\theta < \theta_o|x]\}$$

and by making decision d_2 for those values of x where

$$b\{Pr[\theta \geq \theta_o|x]\} > c\{Pr[\theta < \theta_o|x]\} \quad .$$

Although such a situation is unlikely in practice, when the two possible values of the integral are equal, we will be indifferent about which decision to take.

In many applications of this type, the range of test scores may be cut into a decision d_1 region and a decision d_2 region by considering the posterior distribution of θ[i.e., $p(\theta|x)$] as a function of x . Doing this, we see that the critical point dividing the two regions can be represented by that point x_c in the set of possible test results such that

$$c\{Pr[\theta < \theta_o|x_o]\} = b\{Pr[\theta \geq \theta_o|x_o]\}$$

or

$$c\{Pr[\theta < \theta_o|x_o]\} = b\{1 - Pr[\theta < \theta_o|x_o]\}$$

so that

$$\frac{b}{c + b} = Pr[\theta < \theta_o|x_o] \; .^2 \qquad (2.2)$$

What this equation says is that if we consider the class of possible posterior distributions $\{p(\theta|x)\}$ to be indexed by x, and if we can find that member of this class which is identified by $x = x_o$, say, such that $Pr[\theta < \theta_o|x_o] = b/(c + b)$, then for all $x < x_o$, we will choose decision d_1 and for all $x > x_o$, we will select decision d_2 .

We will illustrate the computation of the cutting score x_o on our observation scale with an example. Suppose that the test score X has a normal distribution with unknown mean θ and <u>known</u> variance σ^2 . Further, suppose that all existing information about the parameter θ, which measures the mastery level of a certain skill, may be summarized by a normal distribution with mean τ and variance ϕ . Then, a simple application of Bayes theorem

[2] The uniqueness of the point x_o satisfying this equation is presumed.

yields a posterior distribution of θ which is normal and has a mean of $(\tau\sigma^2 + x\phi)/(\sigma^2 + \phi)$ and a variance of $\sigma^2\phi/(\phi + \sigma^2)$. Transforming this posterior distribution on θ into standard form, we see that

$$\Pr[\theta < \theta_o | x_o] = \Pr\left[z < \frac{\theta_o - \dfrac{\tau\sigma^2 + x_o\phi}{\sigma^2 + \phi}}{[\phi\sigma^2/(\sigma^2 + \phi)]^{1/2}} \middle| x_o \right] ,$$

where z has a normal distribution with mean zero and variance one. Therefore, the cutting score x_o on the observation scale may be determined by finding that point z_o in the standard normal distribution which has percentile rank $100[b/(c + b)]$, and then solving the equation

$$z_o = \frac{\theta_o - \dfrac{\tau\sigma^2 + x_o\phi}{\sigma^2 + \phi}}{[\phi\sigma^2/(\sigma^2 + \phi)]^{1/2}}$$

for x_o . Thus,

$$x_o = \frac{\phi + \sigma^2}{\phi} \{\theta_o - z_o[\phi\sigma^2/(\phi + \sigma^2)]^{1/2}\} - \frac{\tau\sigma^2}{\phi} \quad . \qquad (2.3)$$

In order to convey some feeling for how this loss structure and the normal data and normal prior distributions interact to produce cutting scores, Table 13 has been provided. In this table, the desired proficiency level $\theta_o = 75$ and the prior mean $\tau = 80$.

As we would expect with ϕ and σ^2 held constant, the cutting score x_o increases with c/b, the relative loss for a false positive as compared with a false negative error. From an intuitive point of view, this relationship makes sense. As c/b increases, false positive errors become relatively more expensive than those of the false negative variety. Consequently, our

Table 13

Loss Constants c	b	Variances Prior ϕ	Data σ^2	Posterior Variance	Posterior Cutting Score x_o	Critical Posterior Mean $\frac{80\sigma^2 + x_o\phi}{\sigma^2 + \phi}$
1	9	9.0	9.0	4.5	64.57	72.28
1	9	9.0	16.0	5.8	57.58	71.93
1	9	9.0	25.0	6.6	48.67	71.71
1	9	16.0	9.0	5.8	67.39	71.93
1	9	16.0	16.0	8.0	62.76	71.38
1	9	16.0	25.0	9.8	56.94	71.00
1	9	25.0	9.0	6.6	68.72	71.71
1	9	25.0	16.0	9.8	65.24	71.00
1	9	25.0	25.0	12.5	60.95	70.47
3	7	9.0	9.0	4.5	67.77	73.89
3	7	9.0	16.0	5.8	62.61	73.74
3	7	9.0	25.0	6.6	56.01	73.65
3	7	16.0	9.0	5.8	70.22	73.74
3	7	16.0	16.0	8.0	67.03	73.52
3	7	16.0	25.0	9.8	62.99	73.36
3	7	25.0	9.0	6.6	71.36	73.65
3	7	25.0	16.0	9.8	69.11	73.36
3	7	25.0	25.0	12.5	66.29	73.14
5	5	9.0	9.0	4.5	70.00	75.00
5	5	9.0	16.0	5.8	66.11	75.00
5	5	9.0	25.0	6.6	61.11	75.00
5	5	16.0	9.0	5.8	72.19	75.00
5	5	16.0	16.0	8.0	70.00	75.00
5	5	16.0	25.0	9.8	67.19	75.00
5	5	25.0	9.0	6.6	73.20	75.00
5	5	25.0	16.0	9.8	71.80	75.00
5	5	25.0	25.0	12.5	70.00	75.00
7	3	9.0	9.0	4.5	72.23	76.11
7	3	9.0	16.0	5.8	69.61	76.26
7	3	9.0	25.0	6.6	66.21	76.35
7	3	16.0	9.0	5.8	74.16	76.26
7	3	16.0	16.0	8.0	72.97	76.48
7	3	16.0	25.0	9.8	71.39	76.64
7	3	25.0	9.0	6.6	75.04	76.35
7	3	25.0	16.0	9.8	74.49	76.64
7	3	25.0	25.0	12.5	73.71	76.86
9	1	9.0	9.0	4.5	75.43	77.72
9	1	9.0	16.0	5.8	74.64	78.07
9	1	9.0	25.0	6.6	73.55	78.29
9	1	16.0	9.0	5.8	76.99	78.07
9	1	16.0	16.0	8.0	77.24	78.62
9	1	16.0	25.0	9.8	77.43	79.00
9	1	25.0	9.0	6.6	77.68	78.29
9	1	25.0	16.0	9.8	78.36	79.00
9	1	25.0	25.0	12.5	79.05	79.53

decision maker decreases his chances of making false positive errors by increasing the cutting score, x_o .

Since the mean of the posterior distribution of θ is a linear function of x, we may reformulate the question of the critical cutting score x_o into a question concerning the critical posterior mean $\mu_o = (\tau\sigma^2 + x_o\phi)/[\phi\sigma^2/(\sigma^2 + \phi)]^{\frac{1}{2}}$ where x_o is determined by Equation (2.3). In this case, we solve the equation

$$z_o = \frac{\theta_o - \mu_o}{[\phi\sigma^2/(\sigma^2 + \phi)]^{\frac{1}{2}}}$$

for μ_o . And analogous to Equation (2.3) above for x_o, we have

$$\mu_o = \theta_o - z_o \, [\phi\sigma^2/(\sigma^2 + \phi)]^{\frac{1}{2}} .$$

In preparing Table 13, we also calculated values of μ_o which appear in the last column. These results can be understood by referring to Figure 2.1. Our procedure says that the critical posterior distribution of θ must be such that 100[b/(c + b)] percent of the probability lies below θ_o . In order to maintain this constant percentage, as the posterior variance increases, the critical posterior mean of θ must necessarily decrease for $\frac{b}{(c + b)} > .5$ and increase for $\frac{b}{(c + b)} < .5$. That is, as the posterior variance of θ increases, the critical posterior mean moves away from θ_o . From an intuitive point of view, this makes sense. It implies that as a decision maker becomes increasingly uncertain about the posterior mean as an estimate of θ, he becomes more cautious, moving his critical mean in the direction of the less costly errors.

The effects of changes in parameter values on x_o is a bit more complicated since we must consider not only the posterior variance, but also the ratio of the prior variance (ϕ) to the variance of sampling

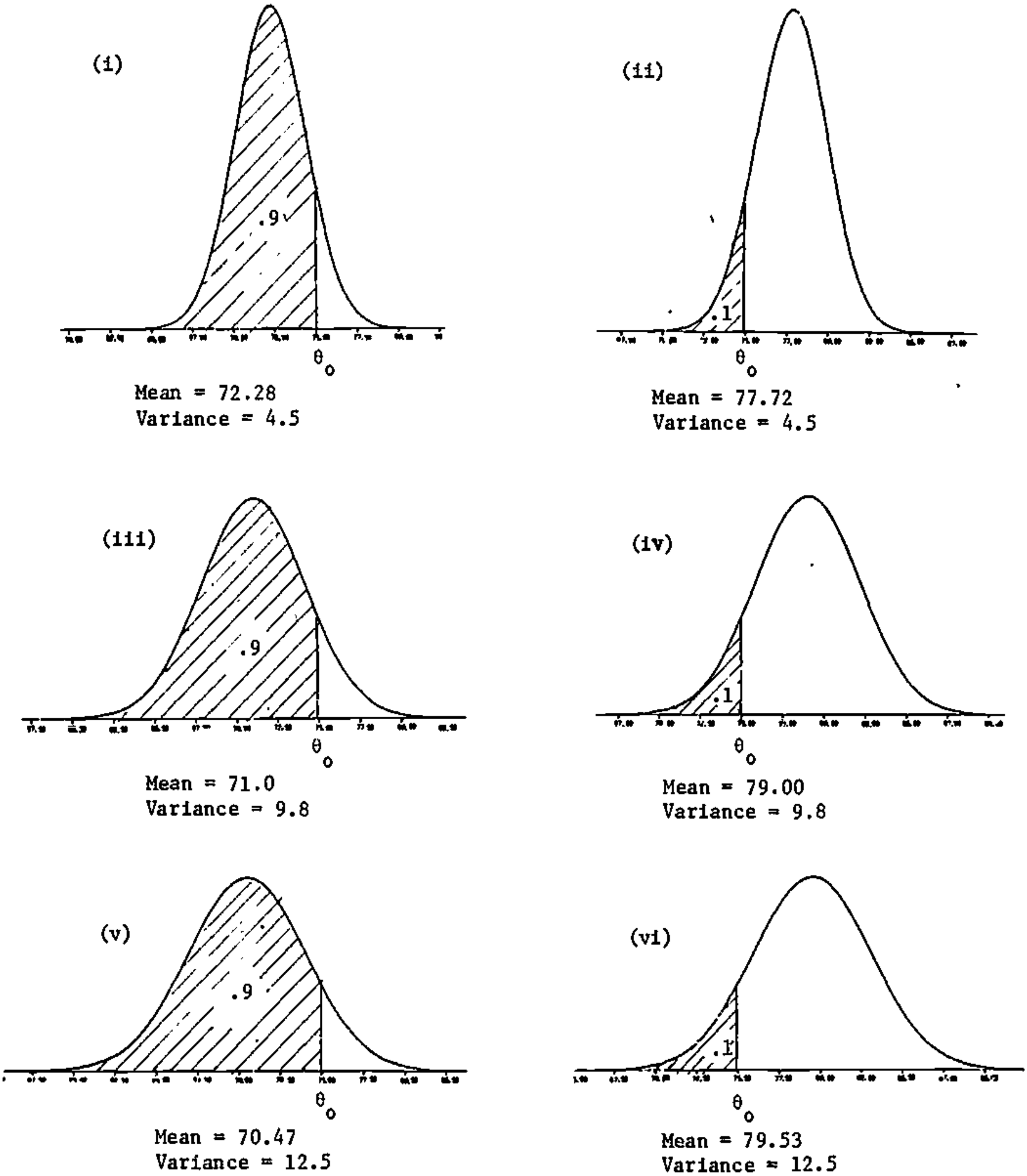


Figure 2.1. This figure illustrates the necessary change in the critical posterior mean as the posterior variance increases. In the figure, we let $b/(b + c) = .9$ in (i), (iii), and (v). We let $b/(b + c) = .1$ in (ii), (iv), and (vi). The figure illustrates that for $b/(b + c) > .5$, the critical posterior mean of θ must decrease with increasing variance, while for $b/(b + c) < .5$ it must increase.

distribution of x (i.e., σ^2). One can get some notion of what is going on by examining the definition of μ_o .

$$\mu_o = \frac{\tau\sigma^2 + x_o\phi}{\sigma^2 + \phi}$$

Since μ_o is the weighted average of τ and x_o, for given μ_o, x_o must increase as the ratio ϕ/σ^2 increases. This relationship is clearly indicated in Table 13.

At this point let us look at a numerical example of some of the theory we have developed in the last few pages. Assume that we have the following situation:

(1) $\rho(\theta) \sim N(80, 25)$ (Prior Distribution)

(2) $p(x|\theta) \sim N(\theta, 16)$ (Distribution of Test Scores)

(3)

$L(d_i, \theta)$ Loss Function

	$\theta < \theta_o$	$\theta \geq \theta_o$
d_1 (retain)	0	1
d_2 (advance)	2	0

(4) $\theta_o = 75$ (Pass Level)

Before collecting our observation on this subject, we can write the posterior probability density function $p(\theta|x)$ as a function of x . Thus,

$$p(\theta|x) \sim N\left[\frac{80(16) + 25x}{16 + 25}, \frac{25(16)}{25 + 16}\right] .$$

Since this is a completely general description of the posterior distribution no matter what value x is observed, it should be possible to determine those values of x which lead us to decision d_1 and those which lead us to d_2. This is what we did symbolically before when we determined the critical score x_o. We defined x_o as that point such that if x happens to fall below x_o, then applying our criterion of minimum expected loss, we would be led to choose d_1. And of course, if x falls above x_o, we would choose d_2. The previous theory tells us that the next step is to calculate the expected loss under each decision. Doing this we find:

$$\mathcal{E}_{\theta|x}[L(d_1, \theta)] = \int_{75}^{\infty} 1\ p(\theta|x)d\theta$$

and

$$\mathcal{E}_{\theta|x}[L(d_2, \theta)] = \int_{-\infty}^{75} 2\ p(\theta|x)d\theta \ .$$

By equating these two expected values symbolically and solving for that value x for which equality obtains, we previously found that

$$x_o = \frac{\phi + \sigma^2}{\phi} \{\theta_o - z_o\ [\phi\sigma^2/(\phi + \sigma^2)]^{\frac{1}{2}}\} - \tau\frac{\sigma^2}{\phi}$$

where z_o is the $(\frac{100b}{c + b})$th percentile of the unit-normal distribution. Substituting into the equation for x_o, we find:

$$x_o = \frac{25 + 16}{25} \{75 - (-.43)\ [25(16)/(25 + 16)]^{\frac{1}{2}}\} - 80\ \frac{16}{25}$$

$$= 74.003 \ .$$

Therefore, we are certain that if we observe an x (test score) which is greater than 74.003, extensive form analysis will lead us to advance the student. Similarly, if the observed test score is less than 74.003, extensive form analysis will cause us to retain the student for additional training. Although, obtaining a score of precisely 74.003 will be somewhat ambiguous

since the two decisions will have the same expected loss, in practice this will not be a problem. We will seldom find a test score which falls precisely on an indifference point. Suppose that we observe $x = 73$. Since 73 is less than $x_o = 74.003$, we can be certain that extensive form analysis will lead us to retain the student (decision d_1). However, for the sake of those who doubt mathematical arguments, we will calculate the expected losses for comparison.

$$\mathcal{E}[L(d_1, \theta)|x = 73] = \int_{75}^{\infty} 1\ p(\theta|x = 73)d\theta = 1 - \Pr(\theta < 75|x = 73)$$

$$\mathcal{E}[L(d_2, \theta)|x = 73] = \int_{\infty}^{75} 2\ p(\theta|x = 73)d\theta = 2[\Pr(\theta < 75|x = 73)] \ .$$

Since

$$p(\theta|x = 73) \sim N(75.73,\ 9.756),$$

we have

$$\Pr(\theta < 75|x = 73) = .409 \ ,$$

and so

$$\mathcal{E}[L(d_1, \theta)|x = 73] = .591$$

and

$$\mathcal{E}[L(d_2, \theta)|x = 73] = .818 \ .$$

Clearly, we would choose d_1 and retain the student as predicted.

A more realistic model than that just described would recognize the fact that in educational settings the model variance as well as the mean is usually unknown. In most situations, the decision maker will have some information concerning the variation and the region in which the observations will fall; total ignorance would preclude even the proper

choice of a measuring instrument. However, except in very special situations, one's knowledge about the model variance σ^2 is typically not sufficiently precise to warrant the application of the known variance model.

The solution to the inference part of the unknown mean and unknown variance problem has been provided by Novick and Jackson (1974). Beginning with a normal model density with location parameter θ and variance parameter ϕ, and a gamma-normal prior density proportional to

$$b_o(\theta, \phi) \propto \frac{1}{\phi^{\frac{m}{2}+1}} \exp\left\{\frac{m(\theta - w.)^2 + R^2}{-2\phi}\right\},$$

they are able to show that

$$t(\theta) = \frac{[m(m+1)]^{\frac{1}{2}}[\theta - (mw. + x)/(m+1)]}{[R^2 + m(x - w.)^2/(m+1)]^{\frac{1}{2}}} \quad (2.4)$$

has a student's t distribution on m degrees of freedom. In this equation, $R^2/(m + 1)$ is the prior modal estimate of ϕ; w. is the prior modal estimate of θ given that $\phi = R^2/(m + 1)$; and, the parameter m is a weight factor which describes the decision maker's degree of confidence in his estimates. For the details of this development, the reader is referred to Novick and Jackson (1974, Chapter 7).

From Equation (2.4), we see that $t(\theta)$ is linear in θ. Thus, $Pr(\theta < \theta_o|x) = Pr[t(\theta) < t(\theta_o)|x]$. And by using Equation (2.2), we can partition the observation scale into two disjoint regions; one which will lead to decision d_1 and the other which will lead to decision d_2. The process is practically the same as in the known variance case. The only important change is that we now use the t table with m degrees of

freedom instead of a table of the unit normal distribution. Thus, we determine the point t_o which has percentile rank 100b/(c + b) and then solve the equation

$$t_o = \frac{[m(m+1)]^{\frac{1}{2}}[\theta_o - (mw. + x_o)/(m+1)]}{[R^2 + m(x_o - w.)^2/(m+1)]^{\frac{1}{2}}} \tag{2.5}$$

for x_o . Although the process of solving this equation will ordinarily lead to two results, only one of the results will solve Equation (2.5) as stated. The other will be associated with Equation (2.5) with $-t_o$ on the left-hand side.

2.4 An Example with a Binomial Model

Let us examine this same basic problem using a different, and in some ways, a more general model. For pedagogical purposes, this analysis will initially be of the normal form variety, for we will actually exhibit the Bayes risk function. Later we will redo the analysis using the much simpler extensive form. Rather than assuming that the mastery level θ takes on values in an interval, let us return to the original situation where only two values are possible. As before, we use the symbol θ_1 to denote the class of nonmasters and θ_2 the class of masters at any point in time. On the basis of prior information about the student and his training, the decision maker formulates a prior probability distribution on the two-point state space, $\rho(\theta_1) = 1 - p$ and $\rho(\theta_2) = p$, where $0 \leq p \leq 1$. Thus, p represents the "prior" probability that a given student is a master. Clearly, if p were equal to zero or one, no uncertainty would exist and no decision problem would remain. As before, two actions are open to the decision maker: d_1, declare the student a nonmaster and retain him at his present level or, d_2, declare the student a master and advance him to the next level. Analogous to Table 3, our decision maker adopts the familiar threshold loss function:

Table 1.4

$L(d_i, \theta_j)$

	θ_1	θ_2
d_1 (retain)	0	b
d_2 (advance)	c	0

where b and c are nonnegative.

It will be convenient in this and many applications to conceptualize a hypothetical population of tasks for which the mastery judgments are relevant. This done, we define α and β as the conditional probabilities of acceptably completing a randomly selected task from this population, given that the student is a master or nonmaster, respectively. In most applications α will be large and β will be small. Our decision maker plans to construct a mastery test by selecting t tasks from our hypothetical population. For the purposes of this example, we will assume that the tasks are experimentally independent and that the probability of success on each task depends only upon the mastery class to which the student belongs. We denote by X the discrete random variable associated with the number of tasks successfully completed. These assumptions imply that, given θ, the random variable X has a binomial mass function given by

$$p(x|\theta) = \begin{cases} \binom{t}{x}\beta^x(1-\beta)^{t-x} & \text{for } \theta = \theta_1 \\ \binom{t}{x}\alpha^x(1-\alpha)^{t-x} & \text{for } \theta = \theta_2 . \end{cases} \tag{2.6}$$

where $x = 0, 1, 2, \ldots, t$.

If we are to follow the plan of normal form analysis previously outlined, we should now write down all reasonable decision rules. With $t + 1$ possible outcomes (i.e., $x = 0, 1, 2, \ldots, t$), an exhaustive list of decision rules for even moderate t, would be cumbersome, indeed. However, in the opinion of these writers the totality of reasonable decision rules for this problem can be summarized by the relationship

$$\delta_s(x) = \begin{cases} \text{retain if } & x < s \\ \text{advance if } & x \geq s \end{cases} \tag{2.7}$$

for $s = 0, 1, 2, \ldots, t + 1$. What this set of decision rules boils down to is the following: Choosing a decision rule $\delta_s(x)$ is equivalent to choosing a cutting score s on the test score scale such that, if the observed number correct (x) lies below s, the student is retained at the present level, otherwise he is advanced. Selection of the decision rule $\delta_{10}(x)$, for example, would lead one to retain the student if he completes 9 items or less correctly, and advance him, if he obtains a total score of at least 10.

For this problem, the risk $R(\delta_s, \theta_j)$ associated with each (δ_s, θ_j) pair can now be symbolically represented by

$$R(\delta_s, \theta_j) = \mathcal{E}_{x|\theta}\{L[\delta_s(x), \theta_j]\} \quad .$$

Inserting the loss function and model density into this relationship, we find that

$$R(\delta_s, \theta_1) = \sum_{x=s}^{t} c\binom{t}{x}\beta^x(1 - \beta)^{t - x} \tag{2.8a}$$

and

$$R(\delta_s, \theta_2) = \sum_{x=0}^{s-1} b\binom{t}{x}\alpha^x(1 - \alpha)^{t - x}$$

$$= b[1 - \sum_{x=s}^{t} \binom{t}{x}\alpha^x(1 - \alpha)^{t - x}] \ . \tag{2.8b}$$

Thus, the Bayes risk for each δ_s may be represented by

$$r(\delta_s, \rho) = \mathcal{E}_\theta [R(\delta_s, \theta)]$$

$$= (1 - p)c \sum_{x=s}^{t} \binom{t}{x} \beta^x (1 - \beta)^{t - x} + bp[1 - \sum_{x=s}^{t} \binom{t}{x} \alpha^x (1 - \alpha)^{t - x}]$$

$$= bp + \sum_{x=s}^{t} [(1 - p)c\binom{t}{x}\beta^x(1 - \beta)^{t - x} - bp\binom{t}{x}\alpha^x(1 - \alpha)^{t - x}] \; . \tag{2.9}$$

Now it is clear that if the expression inside the summation sign has a structure such that it is positive for all x less than some value x_o and negative for all x greater than x_o, we could minimize the Bayes risk, $r(\delta_s, \rho)$, by choosing $\delta_s(x)$ where s is the smallest integer such that $s \geq x_o$. In fact, this expression has the necessary form. To see this, all one needs to do is set the expression of interest equal to zero and solve

$$c(1 - p)\binom{t}{x_o}\beta^{x_o}(1 - \beta)^{t - x_o} - bp\binom{t}{x_o}\alpha^{x_o}(1 - \alpha)^{t - x_o} = 0$$

for the value(s) of x_o . A few routine manipulations yield the root

$$x_o = \frac{\ln(c/b) + \ln[(1 - p)/p] + t \ln[(1 - \beta)/(1 - c)]}{\ln[(1 - \beta)/(1 - \alpha)] + \ln[\alpha/(\beta)]} \; . \tag{2.10}$$

Since there exists only one root x_o, the expression, thought of as a continuous function of x, can cross the x axis at only one point. Therefore, it must be true that there is exactly one region of the x scale where the expression is positive and exactly one region where it is negative. Although this argument assures us that there is exactly one root, it does not reveal in which region the expression is positive and in which it is negative. We, therefore, re-examine the expression in brackets in Equation (2.9).

For notational convenience, we define Q to be the expression of interest. Thus,

$$Q = (1 - p)c\binom{t}{x}\beta^x(1 - \beta)^{t - x} - bp\binom{t}{x}\alpha^x(1 - \alpha)^{t - x} .$$

Also for convenience, we define a quantity R by the equation

$$R = \frac{(1 - p)c\binom{t}{x}\beta^x(1 - \beta)^{t - x}}{bp\binom{t}{x}\alpha^x(1 - \alpha)^{t - x}}$$

$$= \left(\frac{c}{b}\right) \left(\frac{1 - p}{p}\right) \left(\frac{1 - \beta}{1 - \alpha}\right)^t \left[\frac{\beta(1 - \alpha)}{\alpha(1 - \beta)}\right]^x .$$

Our purpose in the following development is to determine that region in which Q is negative. As a first step, we assume that $Q < 0$ and determine the implications of that assumption. If $Q < 0$, then by merely manipulating the inequality, $Q < 0$, we see that $R < 1$ and, therefore, $\ln R < 0$. Thus,

$$\ln R = \ln\left(\frac{c}{b}\right) + \ln\left(\frac{1 - p}{p}\right) + t \ln\left(\frac{1 - \beta}{1 - \alpha}\right) - x \ln\left[\frac{\alpha(1 - \beta)}{\beta(1 - \alpha)}\right] < 0$$

and so,

$$x \ln\left[\frac{\alpha(1 - \beta)}{\beta(1 - \alpha)}\right] > \ln\left(\frac{c}{b}\right) + \ln\left(\frac{1 - p}{p}\right) + t \ln\left(\frac{1 - \beta}{1 - \alpha}\right) . \quad (2.11)$$

Comparing Equation (2.10) and Equation (2.11), we see that if $\ln\{[\alpha(1 - \beta)]/[\beta(1 - \alpha)]\} > 0$, the condition that $Q < 0$ is satisfied whenever $x > x_o$. And if $\ln\{[\alpha(1 - \beta)]/[\beta(1 - \alpha)]\} < 0$, the condition that $Q < 0$ is satisfied whenever $x < x_o$. But the condition that $\ln\{[\alpha(1 - \beta)]/[\beta(1 - \alpha)]\} < 0$ is satisfied if and only if $[\alpha(1 - \beta)]/[\beta(1 - \alpha)] < 1$, which is equivalent to $\alpha < \beta$. Admitting that the foregoing is a bit confusing, we summarize the results in the following table.

	$x < x_o$	$x > x_o$
$\alpha < \beta$	$Q < 0$	$Q > 0$
$\alpha > \beta$	$Q > 0$	$Q < 0$

The implications of the statements in this table for the decision-making process are two.

(1) If $\alpha > \beta$, then the decision maker will minimize the Bayes risk $r(\delta_s, \rho)$ by choosing $\delta_s(x)$ where s is the smallest integer such that $s \geq x_o$.

(2) If $\alpha < \beta$, the decision function (2.7) is inappropriate. The condition $\alpha < \beta$ implies that the decision maker's model asserts that a nonmaster is more likely to get a particular item correct than a master. Such a model would certainly lead the decision maker to consider a decision function of the form

$$\delta_s(x) = \begin{cases} \text{advance} & \text{if } x < s \\ \text{retain} & \text{if } x \geq s \end{cases}$$

as more appropriate than Equation (2.7).

Since under normal conditions α will be considerably greater than β, implication number one above will usually apply.

Perhaps the most important characteristic of the expression for the critical or cutting test score, Equation (2.10), is its dependence on the probabilities α and β. When α and $1 - \beta$ are both near one, the cutting score will tend to be small. Crudely speaking, in this case it does not take many satisfactory performances to decide whether a student is a master or not. On the other hand as α and β approach one another, it becomes ever more costly to separate the masters and nonmasters. If s

is greater than t, the number of tasks in the test is too small to permit any judgment other than d_1 (retain). If $s \leq 0$, the only feasible decision is to advance the student. This might occur if p, the prior probability of state θ_2, is very close to one. Table 15 gives some examples of critical or cutting scores for some selected parameter values, with the loss function constants $c = 2$ and $b = 1$ in Table 14.

Returning now to familiar ground, we review the implications of the latest wrinkle in our decision making scheme in the context of our initial numerical example. The only structural difference between the present situation and that of Section 2.1 is that we now have t tasks instead of one. For the purposes of this example, let us assume that we have a test of length eight (i.e., $t = 8$). Then, following the procedure outlined for normal form analysis problems, we identify the values of the inputs to the black box.

(a) Specification of Prior Information. From Table 1, we see that our prior beliefs about θ may be summarized by $\rho(\theta_1) = .4$ and $\rho(\theta_2) = .6$. Thus, in the notation of this example, we have $p = .6$, or the odds we would be just willing to give that this student is a master without resort to current test score information are 3/2.

(b) Indicating the Experimental Plan. As we pointed out in the previous development of this example, the plan is for the decision maker to give the student a test composed of 8 tasks. On the basis of this test, the decision maker is to give the student a score which is equal to the number of tasks correctly answered. The assumptions are, of course, that the tasks are equally difficult and experimentally independent, given θ . The distribution of X given θ can therefore be described by the binomial mass function

Table 15

Cutting Scores

(Classify as a master (θ_2) if the number correct equals or exceeds the cutting score)

Prior Prob of master	Number of test tasks	Prob. of success for master (α), Prob of success for nonmaster (β)			
		α β	α β	α β	α β
p	t	.8 .2	.8 .1	.9 .2	.9 .1
.9	10	5	4	6	5
	20	10	8	12	10
	30	15	13	17	15
.7	10	5	5	6	5
	20	10	9	12	10
	30	15	13	18	15
.5	10	6	5	6	6
	20	11	9	12	11
	30	16	13	18	16

$$p(x|\theta) = \begin{cases} \binom{t}{x}\beta^{x}(1-\beta)^{t-x} & \text{for } \theta = \theta_1 \\ \binom{t}{x}\alpha^{x}(1-\alpha)^{t-x} & \text{for } \theta = \theta_2 \end{cases}.$$

From Table 2, we see that for this example $\alpha = .8$ and $\beta = .2$. This states that if our student is a master, the probability that he will give an acceptable response to any task is $\alpha = .8$. If he is a nonmaster, this probability is $\beta = .2$.

(c) Specifying Preferences and Decision Rules. From Table 3, we see that in the notation of the previous theoretical development, our threshold loss may be described by $c = 2$ and $b = 1$. In words, this implies that we would be twice as unhappy (in terms of some measure of loss) if we were to advance a nonmaster than we would be if we retained a master--$L(d_2, \theta_1)/L(d_1, \theta_2) = 2$.

Following the procedure indicated in our general development, we identify 10 reasonable decision functions.

$$\delta_s(x) = \begin{cases} d_1 \text{ (retain)} & x < s \\ d_2 \text{ (advance)} & x \geq s \end{cases}$$

for $s = 0, 1, 2, \ldots, 9$.

The ramifications of these decisions are summarized in the following table.

Decision	Retain	Advance
$\delta_0(x)$	Never	Always
$\delta_1(x)$	if $x = 0$	if $x \geq 1$
$\delta_2(x)$	if $x \leq 1$	if $x \geq 2$
$\delta_3(x)$	if $x \leq 2$	if $x \geq 3$
$\delta_4(x)$	if $x \leq 3$	if $x \geq 4$
$\delta_5(x)$	if $x \leq 4$	if $x \geq 5$
$\delta_6(x)$	if $x \leq 5$	if $x \geq 6$
$\delta_7(x)$	if $x \leq 6$	if $x \geq 7$
$\delta_8(x)$	if $x \leq 7$	if $x \geq 8$
$\delta_9(x)$	Always	Never

Our problem, then, will be to select from these 10 reasonable decisions, that one which will minimize the expected or Bayes risk. Substituting into Equation (2.8a) and (2.8b), we indicate the risk of each (δ_s, θ_j) combination by:

$$R(\delta_s, \theta_1) = 2 \sum_{x=s}^{8} \binom{8}{x} .2^x .8^{8-x}$$

and

$$R(\delta_s, \theta_2) = 1 - \sum_{x=s}^{8} \binom{8}{x} .8^x .2^{8-x} .$$

Thus, from Equation (2.9), we have the Bayes risk given by

$$r(\delta_s, \rho) = .6 + \sum_{x=s}^{8} [.8\binom{8}{x} .2^x .8^{8-x} - .6\binom{8}{x} .8^x .2^{8-x}] . \quad (2.12)$$

Our theory tells us that this function is a minimum if we take s to be the next integer greater than x_o, where x_o is given by:

$$x_o = \frac{\ell n(2) + \ell n(2/3) + 8\ \ell n(4)}{\ell n(4) + \ell n(4)}$$

$$= 4.104\ .$$

And, therefore, the Bayes risk $r(\delta_s, \rho)$ will be a minimum if we choose $\delta_5(x)$. Substituting s = 5 into Equation (2.12), we see that the minimum Bayes risk is $r(\delta_5, \rho) = .042$. To convince the still skeptical reader that δ_5 does, in fact, lead one to the minimum Bayes risk, in Table 16 we have exhibited $r(\delta_s, \rho)$ for each of our decision rules.

Table 16

$r(\delta_s, \rho)$

for c/b = 2, p = .6, and t = 8

s	$r(\delta_s, \rho)$
0	.800
1	.666
2	.397
3	.163
4	.051
5	.042
6	.123
7	.298
8	.499
9	.600

As promised, we shall now reconsider the preceding problem using <u>extensive form analysis</u>. It will be recalled that analyses of the extensive form have two major advantages. First, it is unnecessary to exhibit all reasonable decision rules. And secondly, in minimizing the expected loss, it is necessary to consider only that value of x actually obtained. As we shall see, these simplifications will make this analysis almost trivial.

Recall that the likelihood of X given θ is given by

$$p(x|\theta) = \begin{cases} \binom{t}{x}\beta^x(1-\beta)^{t-x} & \text{for } \theta = \theta_1 \text{ (nonmaster)} \\ \\ \binom{t}{x}\alpha^x(1-\alpha)^{t-x} & \text{for } \theta = \theta_2 \text{ (master)} \end{cases}$$

where $x = 0, 1, 2, \ldots, t$ [see Equation (2.6)], and that $\Pr(\theta = \theta_2) = p$. Combining these two probabilities we see that the joint probability density of θ and X is given by

$$p(x, \theta_j) = \begin{cases} \binom{t}{x}(1-p)\beta^x(1-\beta)^{t-x} & j = 1 \\ \\ \binom{t}{x}p\alpha^x(1-\alpha)^{t-x} & j = 2 \end{cases}$$

where $x = 0, 1, 2, \ldots, t$. Therefore, by Bayes theorem, the posterior distribution of θ is given by

$$p(\theta_j|x) = \begin{cases} \dfrac{(1-p)\beta^x(1-\beta)^{t-x}}{(1-p)\beta^x(1-\beta)^{t-x} + p\alpha^x(1-\alpha)^{t-x}} & \text{for } j = 1 \\ \\ \dfrac{p\alpha^x(1-\alpha)^{t-x}}{(1-p)\beta^x(1-\beta)^{t-x} + p\alpha^x(1-\alpha)^{t-x}} & \text{for } j = 2 \;. \end{cases} \tag{2.13}$$

Combining Table 14 with Equation (2.13), we see that the expected posterior loss under d_1 is given by

$$\mathcal{E}_{\theta|x}\,[L(d_1, \theta)] = \frac{bp\alpha^x(1 - \alpha)^{t - x}}{(1 - p)\beta^x(1 - \beta)^{t - x} + p\alpha^x(1 - \alpha)^{t - x}}$$

and that under d_2 by

$$\mathcal{E}_{\theta|x}\,[L(d_2, \theta)] = \frac{c(1 - p)\beta^x(1 - \beta)^{t - x}}{(1 - p)\beta^x(1 - \beta)^{t - x} + p\alpha^x(1 - \alpha)^{t - x}} \quad .$$

Therefore, extensive form analysis leads one to the decision rule given by

$$\delta^*(x) = \begin{cases} d_1 & \text{if } bp\alpha^x(1 - \alpha)^{t - x} < c(1 - p)\beta^x(1 - \beta)^{t - x} \\ d_2 & \text{if } bp\alpha^x(1 - \alpha)^{t - x} > c(1 - p)\beta^x(1 - \beta)^{t - x} \end{cases}$$

or equivalently,

$$\delta^*(x) = \begin{cases} d_1 & \text{if } x \ln\left[\dfrac{\alpha(1 - \beta)}{\beta(1 - \alpha)}\right] < \ln\left(\dfrac{c}{b}\right) + \ln\left(\dfrac{1 - p}{p}\right) + t \ln\left(\dfrac{1 - \beta}{1 - \alpha}\right) \\ d_2 & \text{if } x \ln\left[\dfrac{\alpha(1 - \beta)}{\beta(1 - \alpha)}\right] > \ln\left(\dfrac{c}{b}\right) + \ln\left(\dfrac{1 - p}{p}\right) + t \ln\left(\dfrac{1 - \beta}{1 - \alpha}\right) \end{cases} \quad .$$

By the argument following Equation (2.11), we see that <u>when $\alpha > \beta$</u> this decision rule is equivalent to that reached under analysis of the normal form. Thus, when $\delta > \beta$,

$$\delta^*(x) = \begin{cases} d_1 & \text{if } x < x_o \\ d_2 & \text{if } x > x_o \end{cases}$$

where

$$x_o = \frac{\ln(c/b) + \ln[(1 - p)/p] + t \ln[(1 - \beta)/(1 - \alpha)]}{\ln[(1 - \beta)/(1 - \alpha)] + \ln[\alpha/\beta]} .$$

Consider again this last experimental situation. A moment's reflection will confirm that increases in the number of tasks in the test will increase the overall loss in the decision problem. Each task included in the test, for example, increases the time the student must devote to testing and involves some commitment of facilities and, perhaps, time on the part of a teacher. If we can assume that a certain fixed cost k is incurred for each task in the test, we can reformulate our decision problem into two parts to take account of this additional loss: The first being the selection of the critical number of items I to be answered correctly, and the second being the selection of the optimal number of tasks T to minimize the total Bayes risk $r[\delta(t)^*, \rho] + tk$ [where $\delta(t)^*$ denotes the optimum decision rule when the test contains t items]. The final term, tk, may be thought of as the cost of testing.

Within the framework of normal form analysis, we would seek a cutting score (s) and a number of test tasks (t) that would minimize the total Bayes risk

$$bp \sum_{x=0}^{s-1} \binom{t}{x} \alpha^x (1 - \alpha)^{t - x} + c(1 - p) \sum_{x=s}^{t} \binom{t}{x} \beta^x (1 - \beta)^{t - x} + tk .$$

The constant unit cost or unit loss k must be on the same scale as the original loss function, if we are to have a valid total Bayes risk.

For each fixed value of t, we already know how to select the critical number of items I to be answered correctly. Consequently, with the aid of a computer, it is easy to determine values of $r[\delta(t)^*, \rho] + tk$ for a range of test lengths t . We can then search the display of values of the

total Bayes risk, searching for T, the optimum level of t . An analogous approach applies within the framework of extensive form analysis.

In our example, with $c = 2$, $b = 1$, $\alpha = .8$, $\beta = .2$, and $p = .6$, suppose the loss associated with administering one item is $k = .01$. This could happen, for example, if we think of each test task as taking up .01 as much time as an instructional unit. If the loss constant $c = 2$ was selected because it is associated with the time loss that will result in repeating two instructional units and $b = 1$ because it is associated with the time loss that will result in repeating one unit, then k will be on the same scale. From Table 17, we can see something of the shape of the total Bayes risk function for this example. As the table clearly indicates, in this case the decision maker would choose a test length of $t = 7$ and a critical test score $I = 4$.

Table 17

Display of Total Bayes Risk

for Various Values of t

No. of Items (t)	Critical Score (I)	$r[\delta(t)^*, \rho] + .01\ t$
0	1	.6000
1	1	.2900
2	2	.2680
3	2	.1756
4	3	.1702
5	3	.1311
6	4	.1329
7	4	.1167
8	5	.1221
9	5	.1174
10	6	.1248
11	6	.1263
12	7	.1348
13	7	.1398
14	8	.1489
15	8	.1559
16	9	.1654
17	9	.1736
18	10	.1833
19	10	.1922
20	11	.2020

$b = 1$; $c = 2$; $\alpha = .8$; $\beta = .2$; $p = .6$; $k = .01$

3. Utility Theory

In this section, we turn our attention to a theory which is much stronger than that illustrated in the previous examples. We will now abandon the notion of loss which we relied upon so heavily in Section 2, in favor of the more generally applicable notion of utility. The major difficulty with using loss functions as previously described, lies in the fact that we have simply assumed their existence. Because of the apparent reliance of the loss function on some scale, be it economic, social, political, or other, it is by no means obvious that such a function should exist.

In marked contrast, utility notions do not require that we invent a different and in some sense arbitrary scaling procedure for each problem we meet. Instead, utility theory uses the notions of ordered personal preferences or desirability of outcomes to scale the consequences of each (d, θ) pair. Although several axiom systems have been proposed to insure the existence of utility functions, these axioms generally require only the very basic relationships between preferences which rationality demands. Although these axioms contain many structural details concerning the nature of outcomes, preferences, and rewards, the most important characteristics of these axioms for applications seem to be the requirements for the comparability of any two outcomes and the coherence of the set of possible comparisons.[3] The first requirement merely assures that for any two outcomes A and B, precisely one of the following situations must obtain:

1) A is preferred to B, or

2) B is preferred to A, or

[3] The interested and mathematically able reader is referred to De Groot (1970, Chapter 7) for a detailed consideration of these axioms.

3) A and B are equally desirable outcomes and, we are, therefore, <u>indifferent</u> as to which occurs.

The second requirement is one of transitivity. It merely asserts that for any three outcomes A, B, and C, if A is preferred to B and B is preferred to C, then it must be that A is preferred to C. The point is that these requirements of comparability and coherence are both simple and reasonable. They are incorporated into our system of preferences without question, for it is generally agreed that any violation of these axioms in practice, if exposed, would be deemed ridiculous and one's system of preferences reconsidered.

As an aside, we note that those readers who like the notion of loss described in the previous section need not despair. As Lindley (1972) points out, in applications it typically seems true that one can define a suitable loss function by

$$L(d, \theta) = \underset{d}{\text{Max}}\left\{u(d, \theta)\right\} - u(d, \theta) \qquad (3.1)$$

if the number of outcomes is finite.

Although we will not devote a great deal of space to the problem of assessing one's utility, we will consider one method which will work in problems where there are a finite number of outcomes.

We can represent the set of outcomes by the following table:

	θ_1	θ_2	θ_3	. . .	θ_n
d_1	C_{11}	C_{12}	C_{13}	. . .	C_{1n}
d_2	C_{21}	C_{22}	C_{23}	. . .	C_{2n}
⋮	⋮	⋮	⋮		⋮
d_m	C_{m1}	C_{m2}	C_{m3}	. . .	C_{mn}

Where C_{22}, say, is the outcome or consequence associated with making decision d_2 when θ_2 is the true state of nature. Let C^* be the most preferred outcome and C_* be the least preferred in the table, and assume that you are given a lottery ticket with a v percent chance of "winning" C^* and a 100 - v percent chance of "winning" C_*. Further, assume that someone has offered to take the ticket off your hands in exchange for C_{ij}. Your task is to discover that value of v such that you would be willing to flip a fair coin to decide between the alternatives:

1) A ticket with a v percent chance on C^* and a 100 - v percent chance on C_*, or

2) Selling your ticket for outcome C_{ij}.

The utility of outcome C_{ij} can then be defined by $u(C_{ij}) = v/100$. This procedure can then be followed for each C_{ij} in turn until utilities have been coherently assigned to each of the outcomes.

We now return to our initial example to illustrate this procedure. Recalling that we have two states, nonmaster and master, and two reasonable decisions, retain and advance, we summarize the outcomes in Table 18.

Table 18

possible outcomes C_{ij}

	θ_1 (nonmaster)	θ_2 (master)
d_1 (retain)	C_{11}	C_{12}
d_2 (advance)	C_{21}	C_{22}

Surely the most desirable outcomes are C_{11} and C_{22}. In either of these cases, we correctly classify the student, so it is probably unreasonable to believe that one should be preferred to the other. Furthermore, on the presumption that if a nonmaster is advanced, he will not only lose the

time required to complete the next unit, but may also become frustrated and discouraged, let us assume that misclassifying a master is much more desirable than misclassifying a nonmaster. So in terms of our notation, we have $C_{11} = C_{22} = C^*$ and $C_{21} = C_*$. In this simplified example, the only remaining problem is to determine $u(C_{12})$. What we need to determine is that value v such that our decision maker would be willing to flip a fair coin to decide which gamble he will take:

1) A lottery which pays off C^* (a correctly classified student) v percent of the time and C_* (a misclassified nonmaster) (100 - v) percent of the time; or
2) A sure C_{12} (misclassified master).

Admittedly, specifying v is not an easy task, but it can be done. In order to accomplish this, our decision maker might be aided by considering "how much better" or more desirable correctly classifying a student is than misclassifying a nonmaster and compare this with how much better correctly classifying a student is than misclassifying a master. If, for example, correctly classifying a student gives you 10 "utiles" more than misclassifying a nonmaster and only 5 "utiles" more than misclassifying a master, then v for C_{12} would be 50 percent. This says that misclassifying a master is half-way between misclassifying a nonmaster and correctly classifying a student on a "utiles" or desirability scale. Assuming that, in fact, v = 50 then $u(C_{12}) = v/100 = .5$. Carrying out the above procedure for C_{11}, C_{21}, and C_{22}, we see that it must be true that $u(C_{11}) = u(C_{22}) = 1$ and $u(C_{21}) = 0$. Summarizing this in Table 19, we have

Table 19

Utilities of (d_i, θ_j)

	θ_1 (nonmaster)	θ_2 (master)
d_1 (retain)	1	.5
d_2 (advance)	0	1

In the next section, we will examine some classes or families of utility functions which may be used to describe a decision maker's preferences in the dichotomous or two-action decision problem. Analogous to the minimization of expected loss in extensive form analysis, decision theory with a utility function requires us to select that decision which will <u>maximize</u> the posterior expected or average utility. That is, we seek that $d_i (i = 1, 2)$ such that

$$\int u(d_i, \theta)p(\theta|x)d\theta$$

is a maximum. As usual, θ is the parameter which summarizes the state of nature and $p(\theta|x)$ is its posterior density.

In what follows, it will be important to recognize that if

$$\mathcal{E}_{\theta|x}[u(d_1, \theta)] > \mathcal{E}_{\theta|x}[u(d_2, \theta)] ,$$

then for $b > 0$

$$\mathcal{E}_{\theta|x}[bu(d_1, \theta) + c] > \mathcal{E}_{\theta|x}[bu(d_2, \theta) + c] .$$

That is, if d_1 is preferred to d_2 using the utility function $u(d_i, \theta)$, then d_1 will still be preferred using any positive linear transformation of $u(d_i, \theta)$. A similar demonstration is valid when the direction of

the inequalities is reversed. So positive linear transformations of utility functions will not alter the ultimate decision. This means that for decision purposes, a utility function needs to be determined only up to a positive multiplicative constant and an additive constant.

4. Utility and the Two Action Problem

As we saw in the last section, utility functions are imprecise things. In practical situations, we are often unable to specify our utility associated with various (d_i, θ) pairs with anything approaching mathematical precision. Generally, the best we can do is find some approximation which agrees fairly well with our subjective evaluation of the payoffs.

Our purpose in this section will be to illustrate a variety of families or sets of utility functions which have proven useful in applications. In selecting families of utility functions for inclusion in this section, we have used two principal criteria. First of all, we have sought to include families which are mathematically tractable in the sense that their expected values are easily calculated for standard distributions. Secondly, we have sought families which permit an acceptable compromise between having too many parameters for the decision maker to conveniently specify, and being so restricted that significant aspects of the decision maker's preferences cannot be expressed.

4.1 Threshold Utility

With threshold utility, like threshold loss discussed in Section 2, we separate (or partition) the possible values of our state parameter θ into a number of mutually exclusive subsets. For continuous θ we might consider the partition $\{A_1, A_2\}$ where $A_1 = \{\theta | \theta < \theta_o\}$ and $A_2 = \{\theta | \theta \geq \theta_o\}$ for some θ_o . Thus, for decision purposes, those values of θ which are less thar ~ome point, θ_o, will be considered as a set and denoted A_1 . Those values of θ which are greater than or equal to θ_o will be grouped together as A_2 . This partition would be analogous to our previous

examples where θ was a measure of student ability and θ_o was that point which separated the masters from the nonmasters on an educational test.

Another possibility with threshold utility would be to use a slightly finer partition of θ . This might be accomplished, for example, by preselecting three points instead of one: θ_1, θ_2, and θ_3, say. We could then use the partition $\{B_1, B_2, B_3, B_4\}$ where $B_1 = \{\theta|\theta < \theta_1\}$; $B_2 = \{\theta|\theta_1 \leq \theta < \theta_2\}$; $B_3 = \{\theta|\theta_2 \leq \theta < \theta_3\}$; and $B_4 = \{\theta|\theta \geq \theta_3\}$. Such a scheme might be useful if we needed to differentiate those masters who "just barely made the grade" from those who had truly assimilated the material, with analogous distinctions for the nonmasters. Naturally, even finer partitions could be used if the situation warranted it.

Returning to our dichotomous partition $\{A_1, A_2\}$, let us consider it in more detail. If we denote the utility associated with each (d_i, A_j) pair by $u(d_i, A_j)$, then in the two action problem, we may represent the threshold utility function as in Table 20.

Table 20

$u(d_i, A_j)$

	$A_1(\theta < \theta_o)$	$A_2(\theta \geq \theta_o)$
d_1	a	b
d_2	c	d

Decision or action d_1 is then to be preferred whenever

$$\mathcal{E}_{\theta|x}[u(d_1, A_j)] > \mathcal{E}_{\theta|x}[u(d_2, A_j)] ,$$

or whenever

$$a\{Pr[\theta < \theta_o]\} + b\{Pr[\theta \geq \theta_o]\} > c\{Pr[\theta < \theta_o]\} + d\{Pr[\theta \geq \theta_o]\} . \quad (4.1)$$

In decision problems similar to those discussed earlier in the context of educational testing, the d's involve decisions about the true state of nature θ: "Is he a master or a nonmaster?" When this is the case, it will usually happen that either a and d or b and c will be associated with "correct" decisions, and therefore, ought to be larger than the other utilities in their respective columns in the table. Since the labeling of d_1 and d_2 is arbitrary, we will assume that a and d are the correct decisions in this analysis. Furthermore, since we demonstrated at the end of the last section that rescaling utility by a positive linear transformation does not affect our decisions, we will assume that all utilities in the table lie in the interval between zero and one. Applying these stipulations to Equation (4.1), we see that decision d_1 will be preferred whenever

$$(a - c)\{Pr[\theta < \theta_o]\} > (d - b)\{Pr[\theta \geq \theta_o]\}$$

or alternatively,

$$\frac{d - b}{a - c} < \frac{Pr[\theta < \theta_o]}{Pr[\theta \geq \theta_o]} \quad . \tag{4.2}$$

Since

$$Pr[\theta \geq \theta_o] = 1 - Pr[\theta < \theta_o] \ ,$$

we ha e

$$\frac{d - b}{(a - c) + (d - b)} < Pr[\theta < \theta_o] \ . \tag{4.3}$$

By a similar argument, d_2 will be preferred whenever

$$\frac{d - b}{(a - c) + (d - b)} > Pr[\theta < \theta_o] \ .$$

The reader should note the similarity between Equation (4.3) and Equation (2.2). In fact, if a = d = 1 and one considers the loss as specified in Equation (3.1), then it is clear that the equations are identical. And so by simply rephrasing the example of Section 2.3 in terms of utility instead of loss, we see that the posterior cutting score is given by

$$\frac{1 - b}{2 - b - c} = \Pr[\theta < \theta_o | x_o] \ .$$

Consider the following numerical example previously discussed in Section 2.3. Suppose that the test score X has a normal distribution with <u>unknown</u> mean θ and <u>unknown</u> variance ϕ . Further suppose that after careful consideration of all collateral information available, our decision maker is able to adequately summarize his prior beliefs about θ and ϕ for the student under study in terms of a gamma-normal distribution with parameters m = 9, w. = 80, and R^2 = 144. By Equation (2.4), we see that

$$t(\theta|x) = \frac{[9(10)]^{\frac{1}{2}}[\theta - (9*80 + x)/10]}{[144 + 9(x - 80)^2/10]^{\frac{1}{2}}} \tag{4.4}$$

has a student t distribution on 9 degrees of freedom.

In order to apply Equation (4.3), we need two additional pieces of information: A critical true score θ_o, and a utility function $u(d_1, \theta)$. As before we let θ_o = 75. For purposes of this example, we describe our utility function by the following table.

$u(d_1, \theta)$

	$\theta < \theta_o$	$\theta \geq \theta_o$
d_1 (retain)	.9	.5
d_2 (advance)	0	1

Applying Equation (4.3), we see that we will retain the student if $Pr[\theta < \theta_o] > .5/(.9 + .5) \doteq .36$. And by Equation (4.4), this is equivalent to $Pr[t(\theta|x) < t(75|x)] > .36$, where $t(\theta|x)$ is distributed as a student's t with 9 degrees of freedom. If we have already collected our data so that x is known, this decision criterion can be applied directly. However, if x is not yet known, we can apply the same argument used in Section 2.3 to determine the now familiar cutting score, x_o, which divides the observation scale into two disjoint decision regions. There are two steps in the determination of x_o. First, we must locate the point t_o in a table of the central t distribution with 9 degrees of freedom, which has percentile rank equal to 36. Secondly, we must solve the equation

$$t_o = -.37 = \frac{(90)^{\frac{1}{2}}[75 - (720 + x_o)/10]}{[144 + 9(x_o - 80)^2/10]^{\frac{1}{2}}} \tag{4.5}$$

for x_o. Working out the algebra, we find $x_o = 43.5$ and $x_o = .635$. Substituting these two possible solutions into Equation (4.5), we see that only $x_o = 43.5$ satisfies the equation as stated. Thus, whenever the observed x is greater than 43.5, extensive form analysis will lead the decision maker to advance the student. Of course, whenever x is less than 43.5, the student will be retained.

We turn now to another class of utility functions. This time, we will treat utility as a continuous function of θ rather than a discrete one. Continuous utility might be considered as the limiting case of threshold utility as the partition grows increasingly fine.

4.2 Linear Utility

We direct our attention first to the simplest of continuous utility functions, that is, those which are linear in θ . In the linear two-action case, we define utility by functions of the following form.

$$u(d_i, \theta) = \begin{cases} e + f\theta & i = 1 \\ g + h\theta & i = 2 \end{cases} \qquad (4.6)$$

The reader should note that what we have done here is define utility as a separate linear function for each possible decision, d_i . Thus, if decision one is chosen, the payoff or utility is to be a linear function of the state parameter θ with slope f and intercept e . For decision d_2, the slope is h and the intercept is g .

The existence of a breakeven or indifference value θ_o of the state parameter θ imposes the condition that

$$e + f\,\theta_o = g + h\,\theta_o, \text{ or } \theta_o = (e - g)/(h - f) \quad .$$

In our attempt to maximize expected utility, we will select action d_1 if

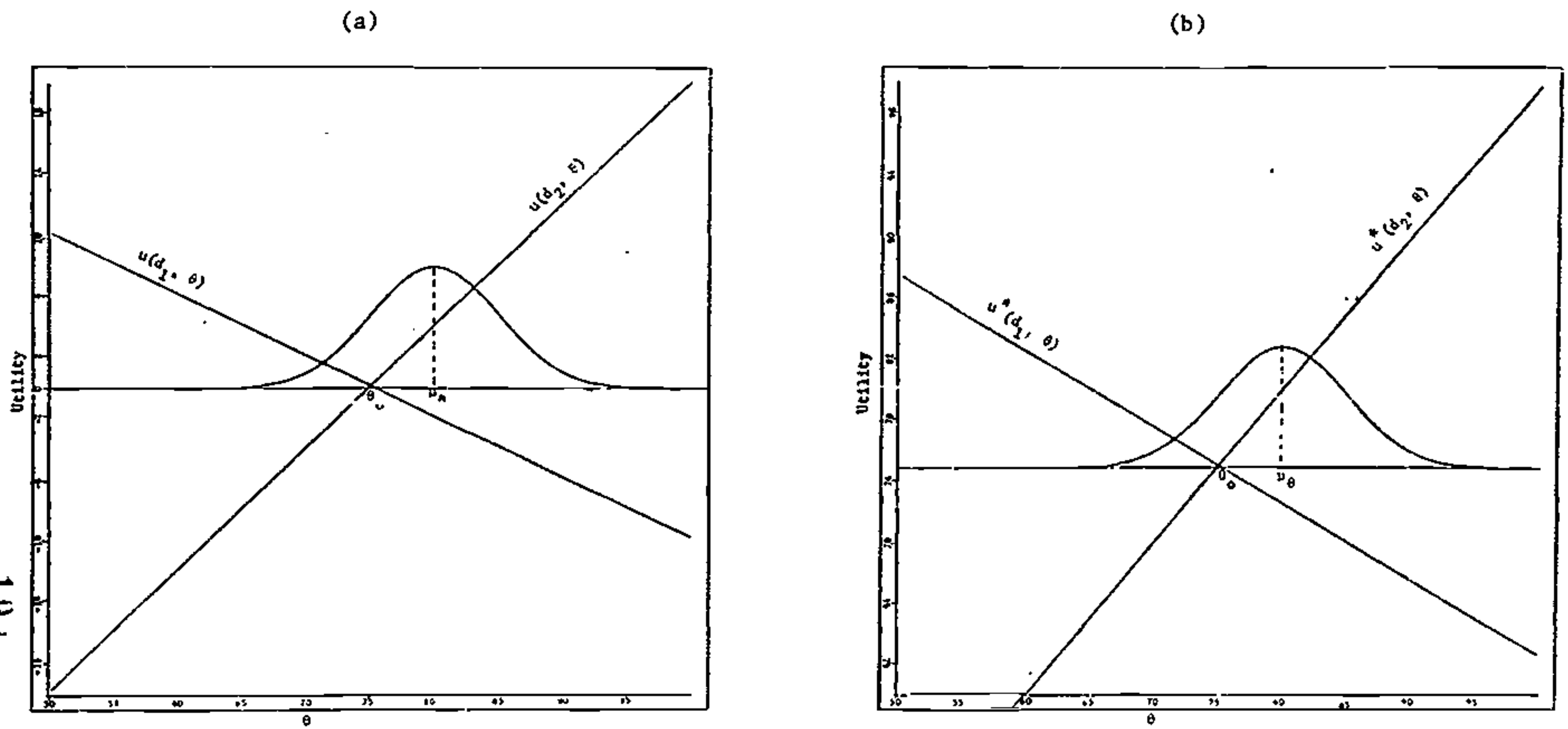
$$\mathcal{E}[(e + f\theta)|x] > \mathcal{E}[(g + h\theta)|x] \quad .$$

If $\mu_\theta = \mu_{\theta|x}$ denotes the posterior mean of θ, this implies that action d_1 is taken if

$$e + f\,\mu_\theta > g + h\,\mu_\theta \; .$$

In other words, with linear utility, the action taken depends only upon the mean of the posterior distribution of the state parameter θ, other attributes of the distribution are irrelevant for decision purposes.

If we index our decisions so that $h - f > 0$, we take action d_1 whenever $\mu_\theta < \theta_o$ and action d_2 whenever $\mu_\theta > \theta_o$. Figure 4.1 illustrates this



<u>Figure 4.1</u>

Linear Utility

Figure (a) above illustrates linear utility of the form of Equation (4.6) with constants e = 30.2, f = -.4, g = -59, and h = .8. Reparameterizing (a) according to $u^*(d_1, \theta)$, Equation (4.7), the utility of θ_o is equated to θ_o and the axes are rotated so that the d_2 branch has a slope equal to one. This reparameterized form of (a) is illustrated in (b). In this illustration, since $u^*(d_2, \mu_\theta) > u^*(d_1, \mu_\theta)$, extensive form analysis will lead the decision maker to select action d_2.

situation graphically. When linear utility is used, one needs to calculate only the utilities of each decision at the posterior mean μ_θ . The decision with the highest utility at μ_θ should then be selected.

As the general linear utility function now stands, Equation (4.6), we need to determine the four constants e, f, g, and h before it is completely specified. However, if we employ the flexibility afforded by the requirement that a utility function needs to be determined only up to a positive linear transformation, we can reduce the number of unknown constants to two. Thus, if $h > 0$ in Equation (4.6), we may redefine $u(d_i, \theta)$ by making the following positive linear transformation

$$u^*(d_i, \theta) = [u(d_i, \theta) - g]/h \ .$$

And so

$$u^*(d_i, \theta) = \begin{cases} e' + f'\theta & i = 1 \\ \\ \theta & i = 2 \end{cases} \tag{4.7}$$

where $e' = (e - g)/h$ and $f' = f/h$.

The nature of our assumption that $h > 0$ for this transformation to be valid cannot be overemphasized. The condition that $h > 0$ is equivalent to the statement that for decision d_2, utility is a strictly increasing function of the state parameter θ . In terms of our previous examples where we considered θ to be an ability index, $h > 0$ would make sense only if d_2 were the decision to advance the student. For if d_2 were the decision to retain him at the present level, we would be in the untenable position of asserting that as ability increases, the utility or desirability of retaining the student at the present level also increases.

If one is careful about making such transformations, this limitation will not cause serious problems. In applications, it is usually the case that the utility or desirability of one of the decisions will increase with the state parameter θ . Thus, all one needs to do is label that decision d_2 and result (4.7) is completely general.

We turn now to an illustration of one of the most direct methods available for determining the constants e' and f' of Equation (4.7). In order to make this method work, the decision maker must be able to specify two ordered pairs (θ_1, θ_2) and (θ_1', θ_2') such that

$$u(d_1, \theta_1) = u(d_2, \theta_2)$$

and

$$u(d_1, \theta_1') = u(d_2, \theta_2') \ .$$

Substituting the equivalents of these expressions from Equation (4.7), we have

$$e' + f'\theta_1 = \theta_2$$

and

$$e' + f'\theta_1' = \theta_2' \ .$$

Solving this system of equations, we find that

$$f' = f/h = \frac{\theta_2 - \theta_2'}{\theta_1 - \theta_1'}$$

and

$$e' = \frac{e - g}{h} = \theta_2 - f'\theta_1 \ .$$

To illustrate the simplicity of using linear utility, consider the following example. After completing a unit of Individually Prescribed Instruction, a student is given a 16 item test to determine whether or not he has mastered the material. From considerable past experience, our decision maker knows that 50% of those students obtaining scores of 12 on the test are able to satisfactorily complete the next sequence. On the basis of this information, he feels that a "true score" of twelve is the minimum necessary for advancing the student to the next unit. Reparameterizing this true score in terms of proportion correct, we find that $\theta_o = .75$.

The next stage in our decision making process consists of determining the posterior distribution on the state parameter θ, where in this problem θ denotes the true proportion correct. Using the techniques described in Novick, Lewis, and Jackson (1973), in Lewis, Wang, and Novick (1973), and in Wang (1973), our decision maker is able to determine a posterior distribution on $\gamma = \sin^{-1}\sqrt{\theta}$. Although this distribution is rather complicated and apparently does not exist in closed form, its precise specification is actually irrelevant for the decision-making process when linear utility is used. Under linear utility, if we can determine or at least approximate the expected value of θ, we will have gleaned all the information from the posterior distribution necessary to make our decision.

Lewis, Wang, and Novick (1973) estimate μ_θ by transforming $\mathcal{E}(\gamma|x)$ according to the equation

$$\text{est of } \mu_\theta = \sin^2[\mathcal{E}(\gamma|x)] .$$

Since this estimate of μ_θ is, in fact, equal to the median of the posterior distribution of θ, it is likely to be a poor estimate only in those cases where the posterior distribution of θ is highly skewed. Furthermore, this is likely to be the case only when the true proportion correct is near either

zero or one. Thus, as long as the critical criterion score θ_o is not too close to either zero or one, errors in estimating μ_θ are unlikely to lead to incorrect advance or retain decisions of major consequence. It is true that if the true proportion correct, θ, is very close to θ_o, an incorrect decision is likely. However, linear utility implies that differences in the utility of the two decisions are not great, for points near the point θ_o. What this means for the decision making process is that with a linear utility function, the output of a readily available and easy to use computer program [see Lewis, Wang, and Novick (1973)] will enable our decision maker to determine a useful estimate of the posterior mean μ_θ at which to evaluate his utility function.

At the next stage in the process, our decision maker must specify his utility function. Actually in this example, very little needs to be done. For most reasonable linear utility functions, the utility associated with the decision to retain the student will have a smaller slope than that associated with the decision to advance the student. Since the two branches of $u(d_i, \theta)$ must intersect at θ_o, he will retain the student if $\mu_\theta < \theta_o$ and advance him if $\mu_\theta \geq \theta_o$. Thus, as long as our decision maker is <u>certain</u> that he will be satisfied with a linear utility function, in the dichotomous decision problem, all he really needs to determine is the ordinal relationship between f and h . If f > h [i.e., the slope of $u(d_1, \theta)$ is greater than the slope of $u(d_2, \theta)$], he will select decision d_1 whenever $\mu_\theta \geq \theta_o$ and select d_2 whenever $\mu_\theta < \theta_o$. Of course if f < h, the situation is reversed.

The "catch" to the foregoing simplicity is that <u>the decision maker is usually not certain that he will be satisfied with a linear utility function until he tries to specify one.</u> In practice the utility function should be overspecified by indicating at least three pairs (θ_i, θ_j) such that

$u(d_i, \theta_i) = u(d_j, \theta_j)$. By overspecification, the decision maker is forced to carefully weigh the implications of a linear utility function.

In applications, linear utility functions seem to behave more reasonably in the neighborhood of the breakeven point, θ_o, than do threshold functions. In this region, the rewards and penalties for correct and incorrect decisions frequently change smoothly rather than abruptly.

The fact that linear utility functions are not bounded when θ is unbounded creates some problems. Several of the axiom systems that have been used to construct decision analysis require that utility functions be bounded. To this theoretical objection must be added the practical fact that unbounded utility functions simply cannot be interpreted far from the breakeven point. These objections are partially removed if the posterior probability distribution of the state parameter θ is fairly closely packed around the breakeven point. If there is almost no probability attached to extreme values, unbounded utility is of little practical importance.

4.3 Quadratic Utility

As we saw above, decisions involving linear utility functions depend only on the mean of the posterior distribution. Quadratic utility functions on the other hand, result in making decisions that depend on both the mean and the variance of the posterior distribution. We begin by defining quadratic utility by a function of the form:

$$u(d_i, \theta) = \begin{cases} -a(\theta - b)(\theta - c) & i = 1 \\ -e(\theta - f)(\theta - g) & i = 2 \end{cases} \quad . \tag{4.8}$$

Observe that in order to use this utility function, we must specify the six constants: a, b, c, e, f, and g . The constants b, c, f, and g have

special meaning in the present parameterization of the utility function. These constants correspond to those values of θ where the utility of the respective decisions is zero. In the example which we considered extensively in Section 2, it might be argued that one reasonable and convenient point which might be used to fix the location of our utility function would be the indifference point θ_o. If we can use the permissible linear transformation of $u(d_i, \theta)$ to force the utility of θ_o equal to zero, we will have established a reference point upon which to judge the utilities associated with other values of θ. In fact, this task can be accomplished by defining

$$u^*(d_1, \theta) = u(d_1, \theta) + a(\theta_o - b)(\theta_o - c)$$

or equivalently

$$u^*(d_1, \theta) = u(d_1, \theta) + e(\theta_o - f)(\theta_o - g) \quad .$$

Recognizing that at the indifference point θ_o, $u(d_1, \theta_o) = u(d_2, \theta_o)$, we can rewrite $u^*(d_i, \theta)$ in the form

$$u^*(d_i, \theta) = \begin{cases} -a(\theta - \theta_o)(\theta - c') & i = 1 \\ -e(\theta - \theta_o)(\theta - g') & i = 2 \end{cases} ;$$

where $c' = b + c - \theta_o$ and $g' = f + g - \theta_o$. Since our permissible linear transformation allows us to specify a scale for utility as well as a location, we may reduce the number of constants to be specified even further by a transformation of the form

$$u^{**}(d_i, \theta) = u^*(d_i, \theta)/a$$

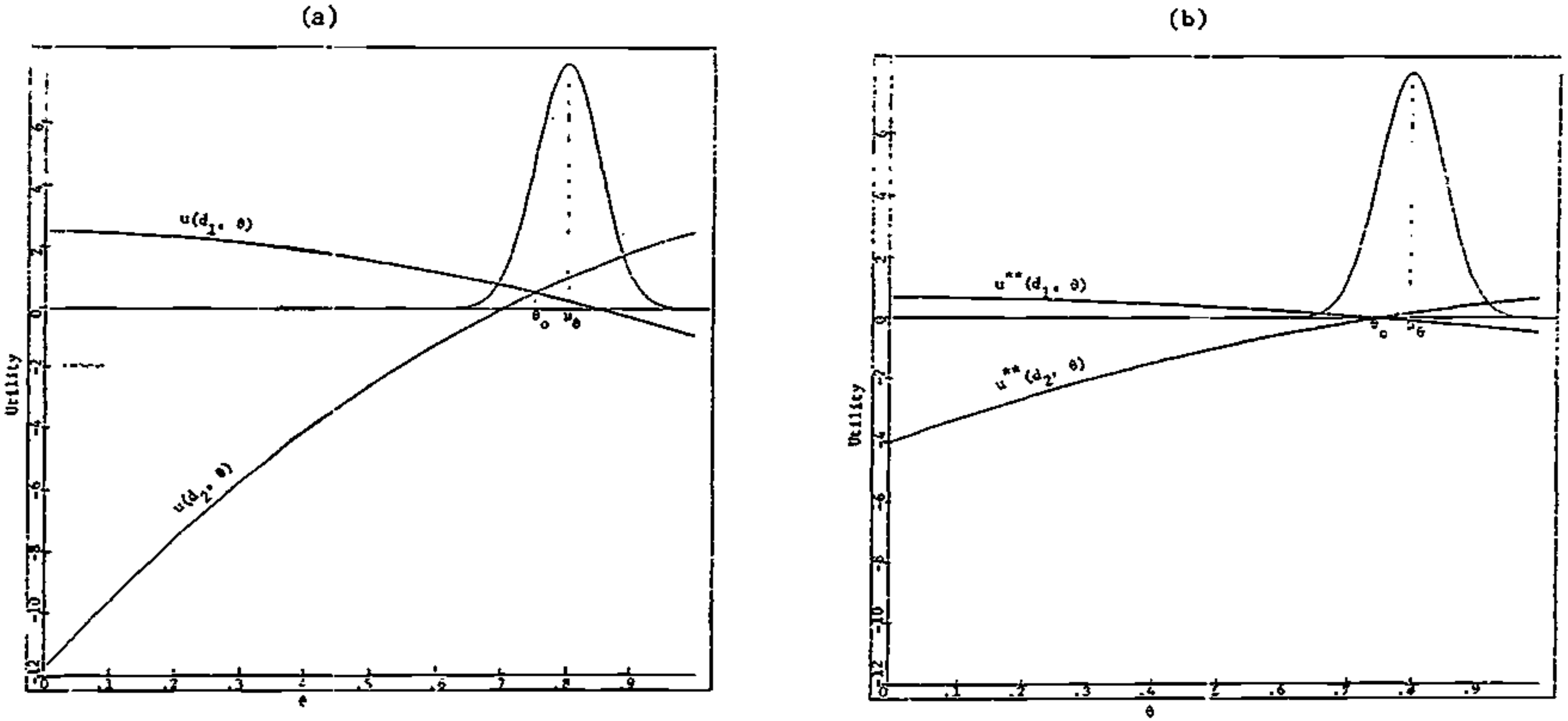
for $a > 0$. Thus,

$$u^{**}(d_i, \theta) = \begin{cases} -(\theta - \theta_o)(\theta - c') & i = 1 \\ -e'(\theta - \theta_o)(\theta - g') & i = 2 \end{cases}$$

where $c' = b + c - \theta_o$, $e' = \frac{e}{a}$, and $g' = f + g - \theta_o$. In order to determine the remaining constants, c', e', and g', the decision maker needs to specify other points on the two branches of the utility function. Although this could be done directly by actually specifying the utilities associated with (d, θ) pairs, we describe a method which is probably easier to use in most situations. This method will frequently work in situations where the utility of each decision seems to approach a maximum asymptotically as the deviation between θ and θ_o increases. In this case, it seems reasonable to situate our quadratic curves so that the convex side is up as illustrated in Figure 4.2. This is equivalent to specifying that both a and e in our original model, Equation (4.8), are positive. We can now fix two of the remaining constants by identifying θ_1 and θ_2 such that θ_1 is at the lower end and θ_2 is at the upper end of the feasible domain of θ . Since we have indicated that the utility approaches a maximum aasymptotically as θ approaches these points, θ_1 and θ_2, it seems reasonable to require that the maximum on the d_1 branch occur at θ_1 and that the maximum on the d_2 branch occur at θ_2 . This requirement is equivalent to the following system of equations:

$$\frac{d}{d\theta} u^{**}(d_1, \theta)\Big|_{\theta = \theta_1} = -(2\theta_1 - \theta_o - c') = 0$$

$$\frac{d}{d\theta} u^{**}(d_2, \theta)\Big|_{\theta = \theta_2} = -e'(2\theta_2 - \theta_o - g') = 0 \quad .$$



<u>Figure 4.2</u>

Quadratic Utility

Figure (a) illustrates a quadratic utility function of the form of Equation (4.8) with constants a = 3, b = −1, c = .85, e = 8.4, f = .7, and g = 2.0. Reparameterizing (a) according to $u^{**}(d_i, \theta)$ leads to (b). Reparameterization $u^{*}(d_i, \theta)$ changes the zero point on the utility scale so that the utility of θ_o is zero. This is equivalent to specifying b = f = θ_o, c = −.9, and g = 1.95 in Equation (4.8). Reparameterization $u^{**}(d_i, \theta)$ then alters the scale of utility so that a = 1, forcing e to equal 2.8.

Thus, $c' = 2\theta_1 - \theta_o$ and $g' = 2\theta_2 - \theta_o$. This procedure assumes, as is true in our examples, that large values of θ make action d_2 more desirable and that small values of θ make d_1 more favorable. Since the indexing of the decisions is arbitrary, however, this restriction is not serious. The remaining constant e' may now be determined by specifying a pair of state parameters (θ_3, θ_4) such that $u(d_1, \theta_3) = u(d_2, \theta_4)$. Since e' is the only unknown in this equation, it can be easily determined.

To illustrate these computations, let us reconsider the example used in the previous section with linear utility. In this example, the point at which the decision maker would be indifferent whether he advanced (d_2) or retained (d_1) the student was $\theta_o = .75$. Since θ is the "true" proportion correct, the minimum feasible θ is $\theta_1 = 0$ and the maximum is $\theta_2 = 1.0$. Thus, solving c' and g' in the equations above, we find $c' = -.75$ and $g' = 1.25$. If, in addition, the decision maker feels that $u(d_1, .7) = u(d_2, .85)$, say, then e' can be found by solving the equation

$$(.7 - .75)(.7 + .75) = e'(.85 - .75)(.85 - 1.25)$$

or

$$e' = 1.8 .$$

The utility function is illustrated in Figure 4.3. In general, the final decision will be for action d_1 if

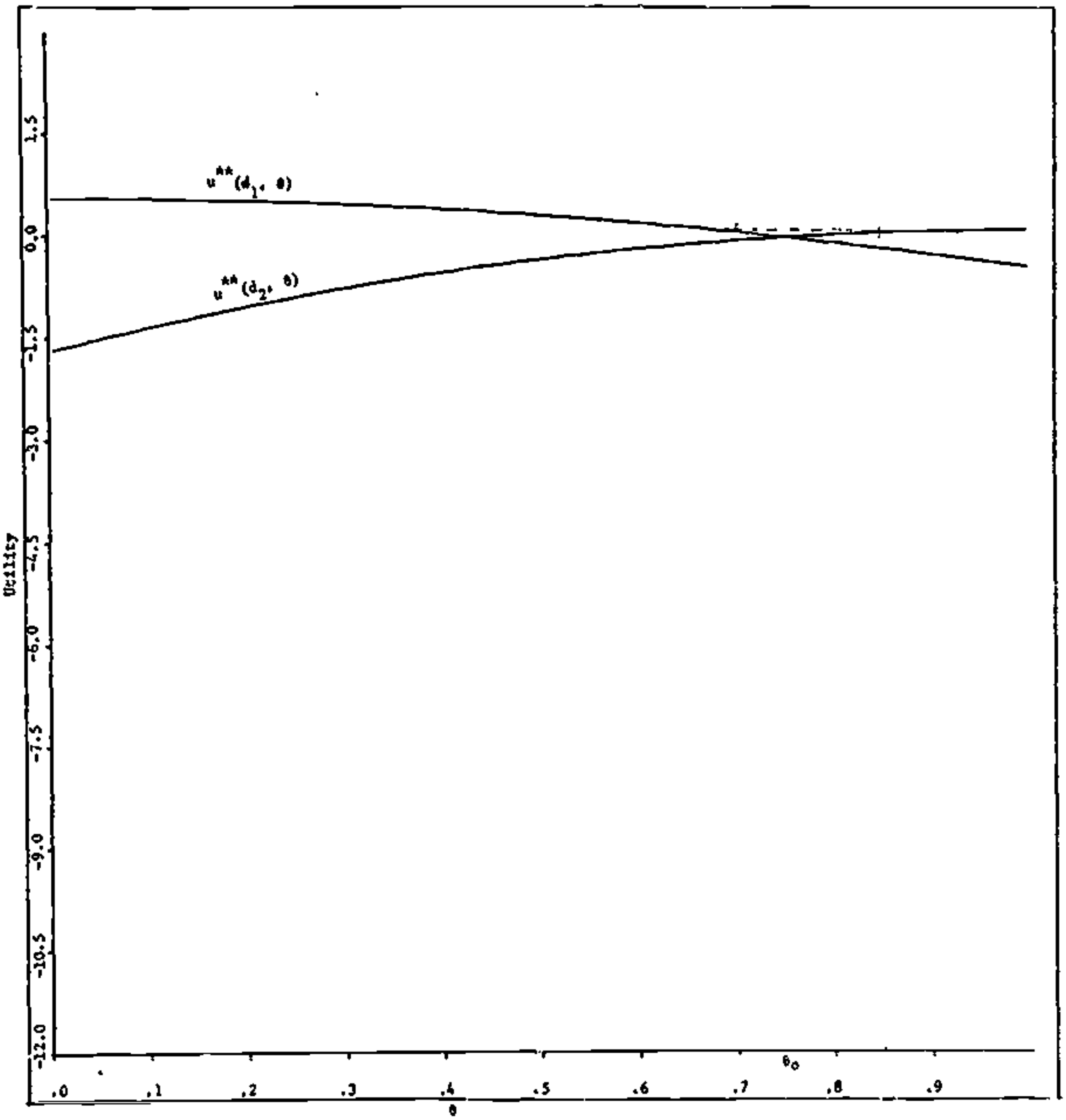
$$\mathcal{E}[u^{**}(d_1, \theta)] > \mathcal{E}[u^{**}(d_2, \theta)]$$

where the expectation is taken with respect to the posterior distribution of θ. Thus, d_1 is to be preferred whenever

$$\sigma_\theta^2 + (\mu_\theta - \theta_o)(\mu_\theta - c') < e'\sigma_\theta^2 + e'(\mu_\theta - \theta_o)(\mu_\theta - g')$$

or

$$u^{**}(d_1, \mu_\theta) - \sigma_\theta^2 > u^{**}(d_2, \mu_\theta) - e'\sigma_\theta^2 ,$$



<u>Figure 4.3</u>

Quadratic Utility

This figure illustrates quadratic utility as transformed by $u^{**}(d_1, \theta)$. In this example, $a' = -.75$, $g' = 1.25$, $e' = 1.8$, $\theta_o = .75$, $\theta_1 = 0.0$, and $\theta_2 = 1.0$. Observe that $u^{**}(d_1, \theta)$ approaches its maximum at zero while $u^{**}(d_2, \theta)$ approaches its maximum at one. Also note that $u(d_1, .7) = u(d_2, .85)$.

where μ_θ and σ_θ^2 are the mean and variance of the posterior distribution of the state parameter θ . This result may be interpreted as calling for action d_1 whenever the utility of d_1 at μ_θ is greater than the utility of d_2 at μ_θ plus a correction factor. The correction factor depends upon the posterior variance of θ and consequently is a measure of the probable deviation of θ from μ_θ . For our Individually Prescribed Instruction examples, this relationship has important consequences. It indicates that for decision purposes, all we need from the posterior distribution of θ is its mean and its variance. In applications, we can often readily obtain these values, or approximations to them, even when the posterior distribution of θ does not exist in closed form.

4.4 Exponential Utility

Linear and quadratic utility functions have played important roles in applications of decision analysis. As we have seen, linear utility requires only that we evaluate the posterior mean, while quadratic utility requires both the posterior mean and variance. These simplifications of the decision process are extremely important, especially when the posterior distribution of θ is of a complicated form. Often we are able to estimate the mean and sometimes the variance of θ, even when the posterior density itself does not exist in closed form. As we shall see in this section, exponential utility also has a simplifying property which makes it particularly useful with many of the standard posterior density functions. Before illustrating this special property of exponential utility, we will exhibit the form of the function and perform our usual simplifications. We define exponential utility by a function of the following form:

$$u(d_i, \theta) = \begin{cases} c - a * \exp\{b\theta\} & i = 1 \\ \\ c - f * \exp\{-g\theta\} & i = 2 \end{cases} \qquad (4.9a)$$

where the constants a, b, f, and g are positive. Notice that this is not the most general form available, since we require that a, b, f, and g be positive and that the <u>same</u> constant c appears under each decision rule. Although this simplification is made so that the estimates of the constants are more easily obtained, it has certain implications which the would be user should keep in mind. This particular formulation requires that for decision d_1, utility is a decreasing function of the state parameter θ. On the other hand, for decision d_2, utility must increase with increasing θ. To see this, all we need to do is rewrite $u(d_2, \theta)$ in the form $u(d_2, \theta) = c - f/\exp\{g\theta\}$. Clearly as θ increases, $f/\exp\{g\theta\}$ approaches zero. That is, as θ increases, the contribution of the second term to utility decreases rapidly, with $u(d_2, \theta)$ approaching c from below as an asymptote. A similar argument with respect to $u(d_1, \theta)$ shows that the utility of decision d_1 also approaches c from below, but this time with <u>decreasing</u> θ. This relationship is depicted graphically in Figure 4.4. Using our now familiar permissible linear transformation to eliminate some of the unknown constants, we let

$$u^*(d_i, \theta) = \frac{1}{a * \exp\{b\theta_o\}} \, [u(d_i, \theta) - (c - a * \exp\{b\theta_o\})] \,.$$

Since $a * \exp\{b\theta_o\} = f * \exp\{-g\theta_o\}$, we have

$$u^*(d_i, \theta) = \begin{cases} 1 - \exp\{b(\theta - \theta_o)\} & i = 1 \\ \\ 1 - \exp\{-g(\theta - \theta_o)\} & i = 2 \end{cases} \,. \qquad (4.9b)$$

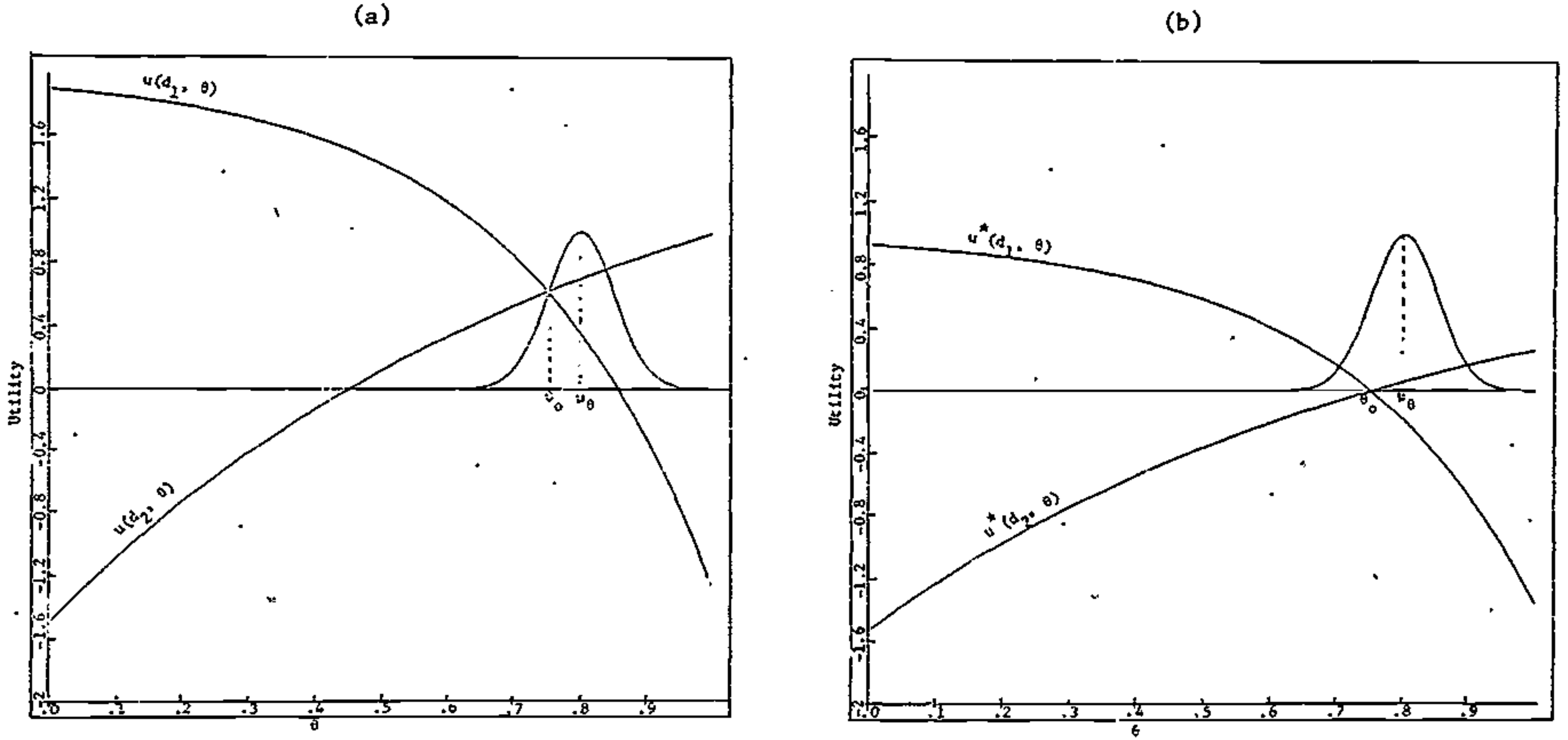


Figure 4.4

Exponential Utility

Figure (a) illustrates $u(d_i, \theta)$ from Equation (4.9) with $a = 0.1$, $b = 3.5$, $c = 2.0$, $f = 3.5$, and $g = 1.24$. Figure (b) is $u^*(d_i, \theta)$, the reparameterized form of figure (a). Upon reparameterization, the utility of θ_o is zero and the scale is changed so that $c = a = f = 1.0$. In its reparameterized form, $u^*(d_i, \theta)$, an exponential utility function is conpletely determined once θ_o and the slopes at θ_o, $-b$ and g, of its two branches are specified.

When the decision maker turns to fixing the two parameters of the utility function, b and g, he must specify the <u>precise</u> utility of at least one point on each branch of the utility function in addition to θ_o . That is, he must specify a pair of points $\{(\theta_1, c_1), (\theta_2, c_2)\}$, such that $u(d_1, \theta_1) = c_1$ and $u(d_2, \theta_2) = c_2$. The reader is cautioned that this is not equivalent to what we have done in the past when we specified points (θ_1, θ_2) such that $u(d_1, \theta_1) = u(d_2, \theta_2)$. With exponential utility of the form of Equation (4.9), we must actually specify the values c_1 and c_2 (although they may, of course, be equal). Once this is done, the parameters b and g are completely determined, for we have

$$1 - \exp\{b(\theta_1 - \theta_o)\} = c_1$$

and

$$1 - \exp\{-g(\theta_2 - \theta_o)\} = c_2 .$$

Upon taking logarithms of both sides and solving, we have

$$b = \frac{\ell n\ (1 - c_1)}{(\theta_1 - \theta_o)} \tag{4.10}$$

and

$$g = -\frac{\ell n(1 - c_2)}{(\theta_2 - \theta_o)} \quad . \tag{4.11}$$

From these equations and the restriction on our model that b and g be positive, we see that if $\theta_1 < \theta_o$ and $\theta_2 > \theta_o$, then c_1 and c_2 must lie between zero and one. This is completely reasonable, however. In examining the transformed utility function, we see that it is zero at θ_o and increases to one as θ approaches either $-\infty$ or $+\infty$ for decisions d_1 and d_2, respectively (see Figure 4.4).

Probably the easiest way to specify c_1 in applications is to determine θ_1 such that the utility of (d_1, θ_1) is half way between the maximum possible utility under d_1 and the utility of (d_1, θ_o). By specifying a similar point θ_2 under d_2, the decision maker can check his coherence by observing whether $u(d_1, \theta_1) = u(d_2, \theta_2)$ subjectively. Having established his utility function, the decision maker will prefer decision d_1 whenever

$$\mathcal{E}[1 - \exp\{b(\theta - \theta_o)\} \mid x] > \mathcal{E}[1 - \exp\{-g(\theta - \theta_o)\} \mid x] .$$

But since expectation is a linear operator, this condition is equivalent to preferring d_1 whenever

$$1 - \mathcal{E}[\exp\{b(\theta - \theta_o)\} \mid x] > 1 - \mathcal{E}[\exp\{-g(\theta - \theta_o)\} \mid x]$$

or

$$\exp\{-b\theta_o\} \mathcal{E}[\exp\{b\theta\} \mid x] < \exp\{g\theta_o\} \mathcal{E}[\exp\{-g\theta\} \mid x]$$

or

$$\exp\{-(g + b)\theta_o\} \mathcal{E}[\exp\{b\theta\} \mid x] < \mathcal{E}[\exp\{-g\theta\} \mid x] .$$

But $\mathcal{E}[\exp\{t\theta\}] = M(t)$ has special significance for mathematical statisticians. In the statistical literature, $M(t)$ is referred to as the moment-generating function for θ . And because of the importance of these moment-generating functions, the integration necessary to evaluate the expected value has been worked out for most standard density functions. Therefore, if the posterior distribution of θ is one of the standard densities (Normal, Uniform, Triangular, Gamma, and others), the final stage in the decision-making process is merely a matter of "plugging in" the parameters of the posterior distribution on θ and those of the utility function. By reformulating the decision criterion in terms of moment-generating functions, we see that decision d_1 will be preferred whenever $\exp\{-(g + b)\theta_o\} M(b) < M(-g)$.

Returning to an example considered in Section 2.3, we assume that the posterior distribution of θ is normal with mean $\mu_\theta = (\tau\sigma^2 + x\phi)/(\sigma^2 + \phi)$ and variance $\sigma_\theta^2 = \sigma^2\phi/(\phi + \sigma^2)$. Since the moment-generating function for a normal variable is given by $M(t) = \exp\{t\mu_\theta + t^2\sigma_\theta^2/2\}$, decision d_1 will be preferred whenever

$$\exp\{b\mu_\theta + b^2\sigma_\theta^2/2 - (g + b)\theta_o\} < \exp\{-g\mu_\theta + g^2\sigma_\theta^2/2\}$$

or

$$\mu_\theta < \theta_o + (g - b)\sigma_\theta^2/2 \ . \tag{4.12}$$

That is, the retain decision is preferred if the mean of the posterior distribution of θ is less than θ_o plus an adjustment which depends upon the variance of the posterior distribution and the relative utilities of the two decisions. Whether that adjustment is positive or negative, depends upon the sign of $g - b$. Since b and g are the magnitudes of the slopes of the utility function at θ_o for d_1 and d_2, respectively, the difference is a measure of the relative speed with which utility is changing on its two branches as θ moves away from θ_o . Thus, if $g > b$, the utility of d_2 is changing more rapidly in the vicinity of θ_o than the utility of d_1 . That is, when $g > b$, making a false, positive error will be relatively more expensive than an error of the false, negative variety for equal distances from θ_o . Consequently, the decision maker adjusts his critical point for μ_θ in a positive direction.

Let us reconsider a slight modification of the known variance numerical example presented in Section 2.3. As before, we assume that:

(1) $\rho(\theta) \sim N(80, 25)$ (prior on θ)

(2) $p(x|\theta) \sim N(\theta, 16)$ (likelihood or model density)

(3) $\theta_o = 75$ (critical criterion score).

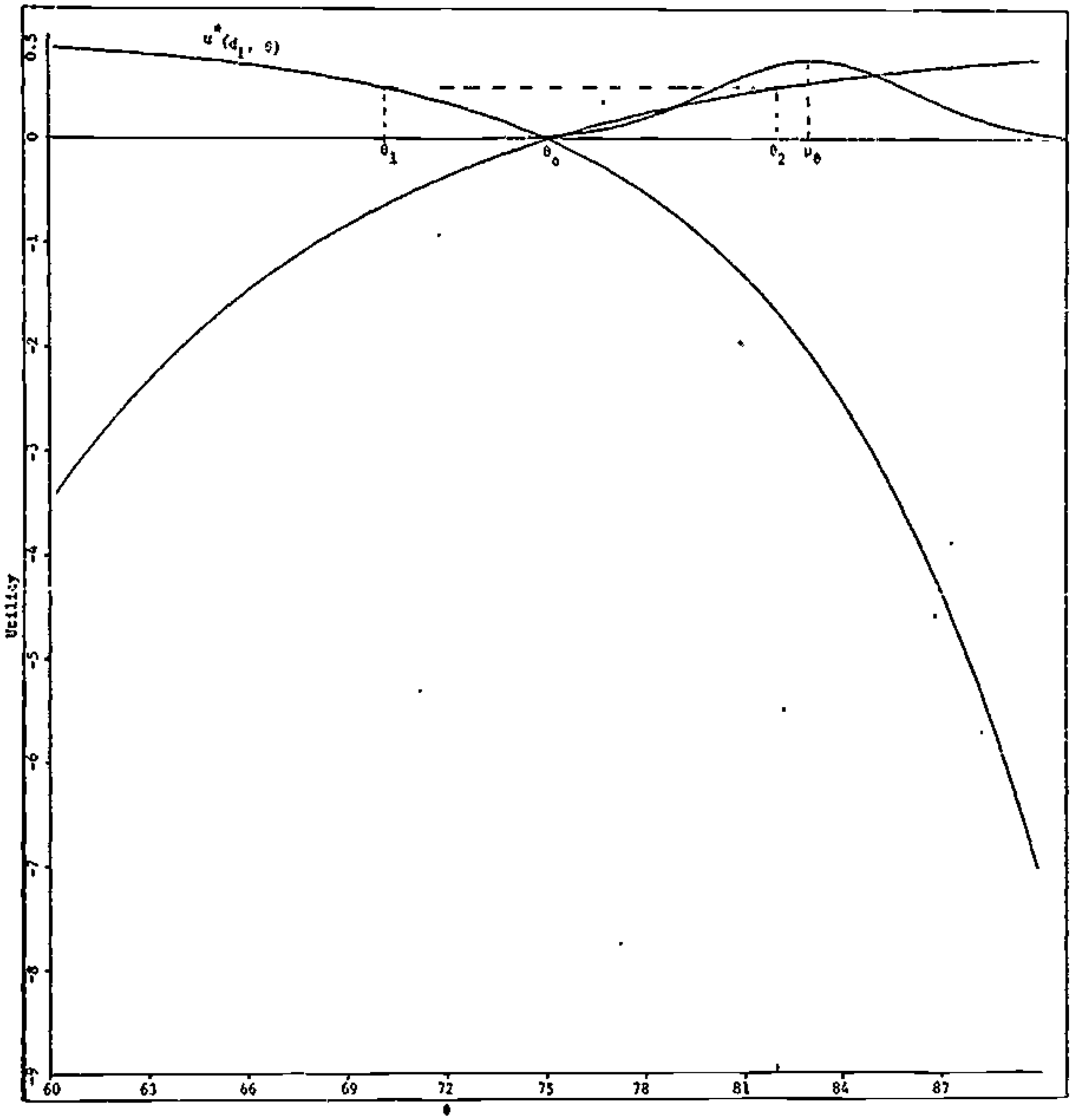


Figure 4.5

Exponential Utility

This figure illustrates exponential utility in the form of Equation (4.9b). In this example, we assumed that $\theta_1 = 70$ and $\theta_2 = 82$. Applying Equations (4.10) and (4.11), we concluded that b is equal to .14 and that g is equal to .10. The posterior distribution of θ is normal with mean equal to 83 and variance equal to 9.76.

We further assume that the student obtains a test score of $x = 85$. Thus, the posterior distribution of θ is normal with mean equal to 83 and variance equal to 9.76.

Next, our decision maker must specify his utility function. On the presumption that he would be satisfied with an exponential description like Equation (4.9), our decision maker would follow the procedures outlined in this section. His first task would be to subjectively determine θ_1 such that the utility of (d_1, θ_1) is half-way between $u(d_1, \theta_o)$ and the maximum possible utility on the d_1 branch. Next he must similarly determine θ_2 . Assuming that $\theta_1 = 70$ and $\theta_2 = 82$ and that these are coherent, Equations (4.10) and (4.11) indicate that $b = .14$ and $g = .10$. Applying Equation (4.12), we see that the student will be retained if $\mu_\theta < [75 + (-.08)9.76/2] = 74.61$. Since $\mu_\theta = 83$, the decision maker will certainly advance the student. This situation is illustrated in Figure 4.5.

4.5 Squared Exponential Utility

The final family of utility functions that we will consider for the two-action problem, will seem somewhat restricted in the amount of flexibility that it permits the decision maker. However, it is very compatible with normal posterior distributions and is frequently useful when other posteriors may be approximated by normal distributions. It is also a rather natural family of utility functions when the problem is an estimation problem and the act and parameter spaces coincide. The model for squared exponential utility is

$$u(d_1, \theta) = \begin{cases} 1 - \exp\left\{\dfrac{-a(\theta - \theta_o)^2}{2}\right\} & i = 1 \text{ and } \theta < \theta_o \\ 1 - \exp\left\{\dfrac{-b(\theta - \theta_o)^2}{2}\right\} & i = 2 \text{ and } \theta > \theta_o \\ 0 & \text{otherwise} \end{cases} \qquad (4.13)$$

where a and b are positive.

Turning to the specification of the parameters a and b, we follow the same procedure as in the case of exponential utility. That is, we specify the utility of at least one point other than θ_o for each branch of the utility function. If these points and their utilities are represented by the pairs $\{(\theta_1, c_1), (\theta_2, c_2)\}$, then

$$1 - \exp\{-\frac{a}{2}(\theta_1 - \theta_o)^2\} = c_1$$

and

$$1 - \exp\{-\frac{b}{2}(\theta_2 - \theta_o)^2\} = c_2 .$$

Upon taking logarithms and solving, we find that

$$a = -\frac{2 \ln(1 - c_1)}{(\theta_1 - \theta_o)^2}$$

and

$$b = -\frac{2 \ln(1 - c_2)}{(\theta_2 - \theta_o)^2} .$$

From these equations and the restriction on our model that a and b are both positive, we see that c_1 and c_2 must lie in the interval (0, 1). Examining Figure 4.6, we see that this is reasonable, for squared-exponential utility as described by Equation (4.13) is bounded between 0 and 1.

As with exponential utility, one way to determine (θ_1, c_1) and (θ_2, c_2) is to look for those points on the d_1 and d_2 branches such that

$$u(d_i, \theta_i) - u(d_i, \theta_o) = \max_{\theta}\{u(d_i, \theta)\} - u(d_i, \theta_i) .$$

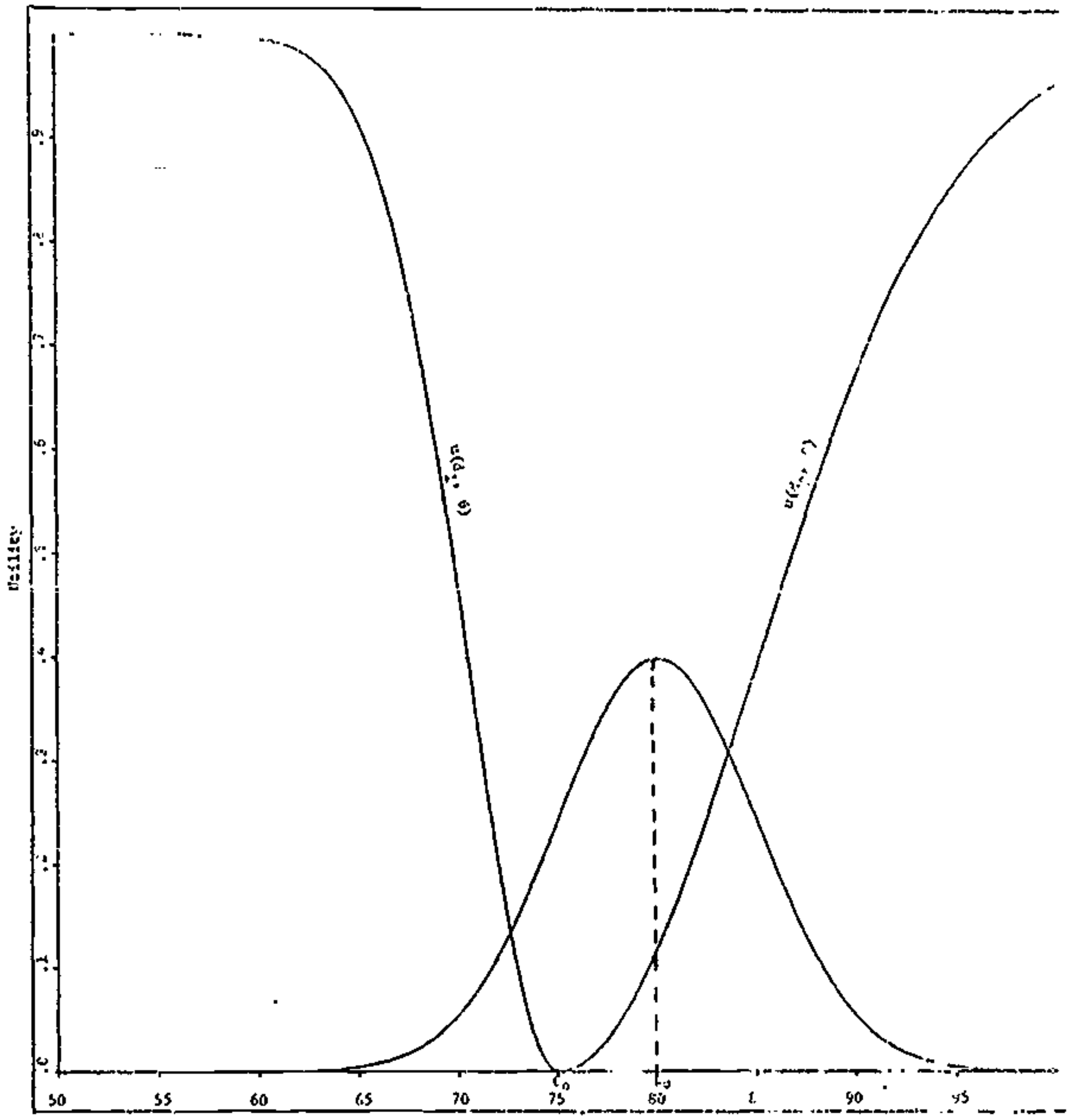


Figure 4.6

Squared Exponential Utility

This figure illustrates squared exponential utility in the form of Equation (4.13). In this figure, a equals .05 and b equals .01. Note that the utility function is bounded between zero and one, with $u(d_1, \theta)$ equal to zero for $\theta \geq \theta_o$ and $u(d_2, \theta)$ equal to zero for $\theta \leq \theta_o$.

That is, try to determine those points (d_1, θ_1) which have desirability approximately half-way between the desirability of (d_1, θ_o) and the maximum possible desirability under that decision. It will then be possible to verify the consistency (coherence) of these specifications by comparing $u(d_1, \theta_1)$ and $u(d_2, \theta_2)$, subjectively. If large differences are believed to exist in the payoff of these two situations, then some reconciliation will be necessary.

Action d_1 will be preferred if the expected or average posterior utility under d_1 is greater than that under d_2. That is, if

$$\int_{-\infty}^{+\infty} u(d_1, \theta)p(\theta|x)d\theta > \int_{-\infty}^{+\infty} u(d_2, \theta)p(\theta|x)d\theta \quad .$$

Since $u(d_1, \theta)$ and $u(d_2, \theta)$ are non-zero only in the regions $\theta < \theta_o$ and $\theta > \theta_o$, respectively, this inequality may be rewritten in the equivalent form

$$\int_{-\infty}^{\theta_o} u(d_1, \theta)p(\theta|x)d\theta > \int_{\theta_o}^{+\infty} u(d_2, \theta)p(\theta|x)d\theta \quad .$$

These integrals will be tractable if the posterior density of θ, $p(\theta|x)$, is normal with mean μ_θ and variance σ_θ^2. In this case,

$$u(d_1, \theta)p(\theta|x) = \frac{1}{\sqrt{2\pi\sigma_\theta^2}}\left[1 - \exp\left\{-\frac{a(\theta - \theta_o)^2}{2}\right\}\right]\left[\exp\left\{-\frac{(\theta - \mu_\theta)^2}{2\sigma_\theta^2}\right\}\right]$$

which may be rewritten in the form

$$u(d_1, \theta)p(\theta|x) = \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left[-\frac{1}{2\sigma_\theta^2}(\theta - \mu_\theta)^2\right]$$

$$- \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left[-\frac{(a + \sigma_\theta^{-2})}{2}\left(\theta - \frac{a\theta_o + \sigma_\theta^{-2}\mu_\theta}{a + \sigma_\theta^{-2}}\right)^2\right] \exp\left[-\tfrac{1}{2}\left(\frac{a\sigma_\theta^{-2}}{a + \sigma_\theta^{-2}}\right)(\theta_o - \mu_\theta)^2\right] .$$

The first term, of course, is nothing but a normal density with mean μ_θ and variance σ_θ^2 . Ignoring that part of the second term which does not depend upon θ, we see that the second term is proportional to a normal distribution with mean $(a\theta_o + \sigma_\theta^{-2}\mu_\theta)/(a + \sigma_\theta^{-2})$ and variance $(a + \sigma_\theta^{-2})^{-1}$. Therefore, $\mathcal{E}[u(d_1, \theta)|x]$ may be written as

$$\mathcal{E}[u(d_1, \theta)|x] = \int_{-\infty}^{\theta_o} \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left\{-\frac{(\theta - \mu_\theta)^2}{2\sigma_\theta^2}\right\} d\theta$$

$$- \frac{1}{\sqrt{\sigma_\theta^2(\sigma_\theta^{-2} + a)}} \exp\left[-\tfrac{1}{2}\left(\frac{a\sigma_\theta^{-2}}{a + \sigma_\theta^{-2}}\right)(\theta_o - \mu_\theta)^2\right]$$

$$* \int_{-\infty}^{\theta_o} \frac{\sqrt{\sigma_\theta^{-2} + a}}{\sqrt{2\pi}} \exp\left[-\left(\frac{a + \sigma_\theta^{-2}}{2}\right)\left(\theta - \frac{a\theta_o + \sigma_\theta^{-2}\mu_\theta}{a + \sigma_\theta^{-2}}\right)^2\right] d\theta .$$

But each of these integrals is nothing but the probability in a tail area f a normal distribution. Therefore, we may express these integrals in terms of the percentile rank of θ_o . Thus, by standardizing θ in each integral so that each distribution is unit normal, we see that

$$100\,\mathcal{E}\left[u(d_1,\ \theta)\,\middle|\,x\right] = \mathrm{PR}\left(\frac{\theta_o - \mu_\theta}{\sqrt{\sigma_\theta^2}}\right)$$

$$- \frac{1}{\sqrt{\sigma_\theta^2(\sigma_\theta^{-2} + a)}}\ \exp\left[-\tfrac{1}{2}\left(\frac{a\sigma_\theta^{-2}}{a + \sigma_\theta^{-2}}\right)(\theta_o - \mu_\theta)^2\right]\ \mathrm{PR}\left(\frac{\theta_o - \dfrac{a\theta_o + \sigma_\theta^{-2}\mu_\theta}{a + \sigma_\theta^{-2}}}{(a + \sigma_\theta^{-2})^{-\frac{1}{2}}}\right)$$

where PR(z) is the percentile rank of z with respect to the unit normal distribution.

Replacing a by b and reversing the limits of integration, we have a similar expression for the posterior expected utility under d_2 . Therefore, our decision criterion may be written in the form: "Choose d_1 if

$$\mathrm{PR}(z_o) - \frac{\mathrm{PR}(z_1)}{\sqrt{\sigma_\theta^2(\sigma_\theta^{-2} + a)}}\exp\left[-\tfrac{1}{2}\left(\frac{\sigma_\theta^{-2}a}{\sigma_\theta^{-2} + a}\right)(\theta_o - \mu_\theta)^2\right]$$

$$> [100 - \mathrm{PR}(z_o)] - \frac{[100 - \mathrm{PR}(z_2)]}{\sqrt{\sigma_\theta^2(\sigma_\theta^{-2} + b)}}\ \exp\left[-\tfrac{1}{2}\left(\frac{\sigma_\theta^{-2}b}{\sigma_\theta^{-2} + b}\right)(\theta_o - \mu_\theta)^2\right],\text{"} \tag{4.14}$$

where

$$z_o = \frac{\theta_o - \mu_\theta}{\sqrt{\sigma_\theta^2}}$$

$$z_1 = \frac{\theta_o - \dfrac{a\theta_o + \sigma_\theta^{-2}\mu_\theta}{a + \sigma_\theta^{-2}}}{(a + \sigma_\theta^{-2})^{-\frac{1}{2}}} = \frac{\sigma_\theta^{-2}}{(a + \sigma_\theta^{-2})^{\frac{1}{2}}}\ (\theta_o - \mu_\theta)$$

$$z_2 = \frac{\theta_o - \dfrac{b\theta_o + \sigma_\theta^{-2}\mu_\theta}{b + \sigma_\theta^{-2}}}{(b + \sigma_\theta^{-2})^{-\frac{1}{2}}} = \frac{\sigma_\theta^{-2}}{(b + \sigma_\theta^{-2})^{\frac{1}{2}}}\ (\theta_o - \mu_\theta)\quad .$$

We prefer d_2 when the direction of the inequality is reversed. Although Equation (4.14) looks rather frightening, it is really rather simple to use once you know the parameters of the utility function and of the posterior distribution on θ. Of course, those decision makers with access to a computer will find its application trivial.

5. A Three Action Example

Section 4 was built around the analysis of a decision problem in which two actions were available to the decision maker. Several families of utility functions were studied and the ideas were illustrated with an example involving the decision to advance or retain a student at a certain level in a sequential chain of instructional steps. Although the notation becomes more complex and the computation a bit more tedious, there are no fundamentally new ideas when we assume that there are three (or any finite number) of options open to the decision maker. In this section, we will illustrate this somewhat more general problem by using natural extensions of two of the families of utility functions discussed earlier.

5.1 Threshold Utility

Consider the following slight modification of the Individually Prescribed Instruction example discussed in Section 2.1. In the previous example, when a student completed a unit of instruction, he was considered a master or a nonmaster and was advanced or retained on the basis of expected utility. In this example, we merely extend the number of levels of mastery by further partitioning the nonmasters into two groups. The first group contains those nonmasters whose ability is close to the cutoff point separating the masters from the nonmasters. The second group contains those who apparently missed the whole point of the lesson. The state of a student being a nonmaster of the poorer variety will be denoted by θ_1; the better nonmasters will be denoted by θ_2; and, the masters by θ_3 .

For purposes of this example, we assume that there are only three actions available to the decision maker. The student may repeat both the present and the previous instructional units; he may repeat only the present unit; or, he may advance to the next unit.

With these specifications, we may now define the usual utility function by the following table.

Table 21

$u(d_i, \theta_j)$

	θ_1	θ_2	θ_3
d_1(back one)	u_{11}	u_{12}	u_{13}
d_2(retain)	u_{21}	u_{22}	u_{23}
d_3(advance)	u_{31}	u_{32}	u_{33}

As we have mentioned before, determination of these utilities is not an easy matter. In Section 3, we described one paradigm for their determination which might be helpful. However, we do not claim that it is the last word in utility specification. Nevertheless, in what follows, we will assume that the decision maker has coherently specified the utilities.

After the test score x is available, the decision will be made by selecting that action d_i, i = 1, 2, 3 which maximizes the posterior expected utility

$$\sum_{j=1}^{3} u(d_i, \theta_j)p(\theta_j|x) \quad .$$

We might think of this problem in terms of specifying two cutting test scores x_o and x_1, where $x_o < x_1$. Then for $x < x_o$, action d_1 will be taken; for $x_o < x < x_1$, action d_2 will be taken; and, for $x > x_1$, action d_3 will be taken. To determine the critical points x_o and x_1 which will divide the range of test scores into a d_1, a d_2, and a d_3 region, we return to a technique described in Section 2. We consider the posterior distribution of θ, $p(\theta|x)$, as a

function of x . Since x_o is the indifference point with respect to decisions d_1 and d_2, at x_o the expected posterior utility under d_1 must equal the expected posterior utility under d_2 . That is,

$$\sum_{j=1}^{3} u_{1j}p(\theta_j|x_o) = \sum_{j=1}^{3} u_{2j}p(\theta_j|x_o) \ .$$

Simplifying this, we see that x_o should be determined so that

$$(u_{11} - u_{13} + u_{23} - u_{21})p(\theta_1|x_o) + (u_{12} - u_{22} - u_{13} + u_{23})p(\theta_2|x_o) + (u_{13} - u_{23}) = 0 \quad . \tag{5.1}$$

Similarly, x_1 should be determined so that

$$(u_{31} - u_{33} + u_{23} - u_{21})p(\theta_1|x_1) + (u_{32} - u_{22} - u_{33} + u_{23})p(\theta_2|x_1) + (u_{33} - u_{23}) = 0 \quad . \tag{5.2}$$

In order to illustrate how to use Equations (5.1) and (5.2) in applications, we return to the example in Section 2.3 where posterior to our observation x, the ability parameter θ was continuous and, in fact, normally distributed. There we described the posterior distribution of θ by

$$p(\theta|x) \sim N\left[\frac{\tau\sigma^2 + x\phi}{\sigma^2 + \phi}, \ \frac{\phi\sigma^2}{\sigma^2 + \phi}\right] \ .$$

For purposes of this example, we also redefine the mastery levels θ_1, θ_2, and θ_3 in terms of critical points T_1 and T_2 on the ability scale (θ). We let $\theta_1 = \{\theta|\theta < T_1\}$, $\theta_2 = \{\theta|T_1 < \theta < T_2\}$, and $\theta_3 = \{\theta|\theta > T_2\}$. Then transforming the posterior $p(\theta|x)$ into a posterior on the normal deviate z, we see that

$$\Pr(\theta_1|x) = \Pr[z < z(T_1, x)],$$

$$\Pr(\theta_2|x) = \Pr[z < z(T_2, x)] - \Pr[z < (T_1, x)],$$

and

$$\Pr(\theta_3|x) = 1 - \Pr[z < z(T_2, x)]$$

where

$$z(T_i, x) = \frac{T_i - \dfrac{\tau\sigma^2 + x\phi}{\sigma^2 + \phi}}{[\phi\sigma^2/(\phi + \sigma^2)]^{\frac{1}{2}}} .$$

So we may rewrite Equations (5.1) and (5.2) in the form

$$(u_{11} + u_{22} - u_{21} - u_{12})\Pr[z < z(T_1, x_o)]$$

$$+ (u_{12} - u_{22} - u_{13} + u_{23})\Pr[z < z(T_2, x_o)]$$

$$+ (u_{13} - u_{23}) = 0$$

and

$$(u_{31} + u_{22} - u_{21} - u_{32})\Pr[z < z(T_1, x_1)]$$

$$+ (u_{32} - u_{22} - u_{33} + u_{23})\Pr[z < z(T_2, x_1)]$$

$$+(u_{33} - u_{23}) = 0 \quad .$$

Each of these equations now needs to be solved iteratively for x_o and x_1 . It is recommended that T_1 and T_2 be used as first approximations to x_o and x_1, respectively.

We now turn to a modification of an example considered in Section 2.3 to illustrate these ideas. Assume that we have the following situation:

(1) $\rho(\theta) \sim N(80, 25)$

(2) $p(x|\theta) \sim N(\theta, 16)$

(3) $T_1 = 60$ and $T_2 = 85$

(4)

$u(d_i, \theta_j)$

	$\theta < T_1$	$T_1 < \theta < T_2$	$\theta > T_2$
d_1 (back one)	7	4	0
d_2 (retain)	2	6	1
d_3 (advance)	1	3	5

Thus, by applying Bayes theorem, the posterior distribution of θ as a function of x may be written in the form:

$$p(\theta|x) \sim N\left[\frac{80(16) + 25x}{16 + 25}, \frac{25(16)}{25 + 16}\right] .$$

Substituting into our equations for $z(T_i, x)$, we have

$$z(T_1, x) = \frac{60 - \dfrac{1280 + 25x}{41}}{3.123} = 9.214 - .195x$$

and

$$z(T_2, x) = \frac{85 - \dfrac{1280 + 25x}{41}}{3.123} = 17.218 - .195x .$$

And we must solve the equations

$$.07 \text{ PR}(9.214 - .195x_o) - .01 \text{ PR}(17.218 - .195x_o) - 1 = 0$$

$$.02 \text{ PR}(9.214 - .195x_1) - .07 \text{ PR}(17.218 - .195x_1) + 4 = 0 ,$$

where PR(z) equals the percentile rank of z . Iterating to a solution, we find that $x_o = 52.7$ and $x_1 = 87.4$. So the decision maker will choose decision d_1 and have the student repeat two units if $x \leq 52$; will choose decision d_1 and have the student repeat the current unit if $53 \leq x \leq 87$; and, will choose decision d_3 and advance the student if $x \geq 88$.

5.2 Linear Utility

Analogous to the situation in the two-action problem (see Section 4.2), we define linear utility to be linear in θ for each decision separately. Thus, linear utility in the three decision situation is defined by a function of the form:

$$u(d_i, \theta) = \begin{cases} e + f\theta & i = 1 \\ g + h\theta & i = 2 \\ k + m\theta & i = 3 \end{cases} \quad . \qquad (5.3)$$

If we assume that our decisions can be indexed so that decision d_1 is most desirable when θ is small, so that decision d_2 is most desirable when θ takes intermediate values, and so that decision d_3 is most desirable when θ is large; then the solution of the three action problem is a straightforward extension of that offered in Section 4.2. Applying our permissible positive linear transformation, for m > 0, we let $u^*(d_i, \theta) = [u(d_i, \theta) - k]/m$. Thus,

$$u^*(d_i, \theta) = \begin{cases} e' + f'\theta & i = 1 \\ g' + h'\theta & i = 2 \\ \theta & i = 3 \end{cases}$$

where

$$e' = \frac{e - k}{m} \; ; \; f' = \frac{f}{m} \; ; \; g' = \frac{g - k}{m} \; ; \text{ and } h' = \frac{h}{m} \quad .$$

And since we have four constants to estimate in specifying our linear utility function, we need four ordered pairs (θ_i, θ_j) such that $u^*(d_i, \theta_i) = u^*(d_j, \theta_j)$. Two of these pairs are provided by the breakeven points T_1 and T_2. At these points, we have

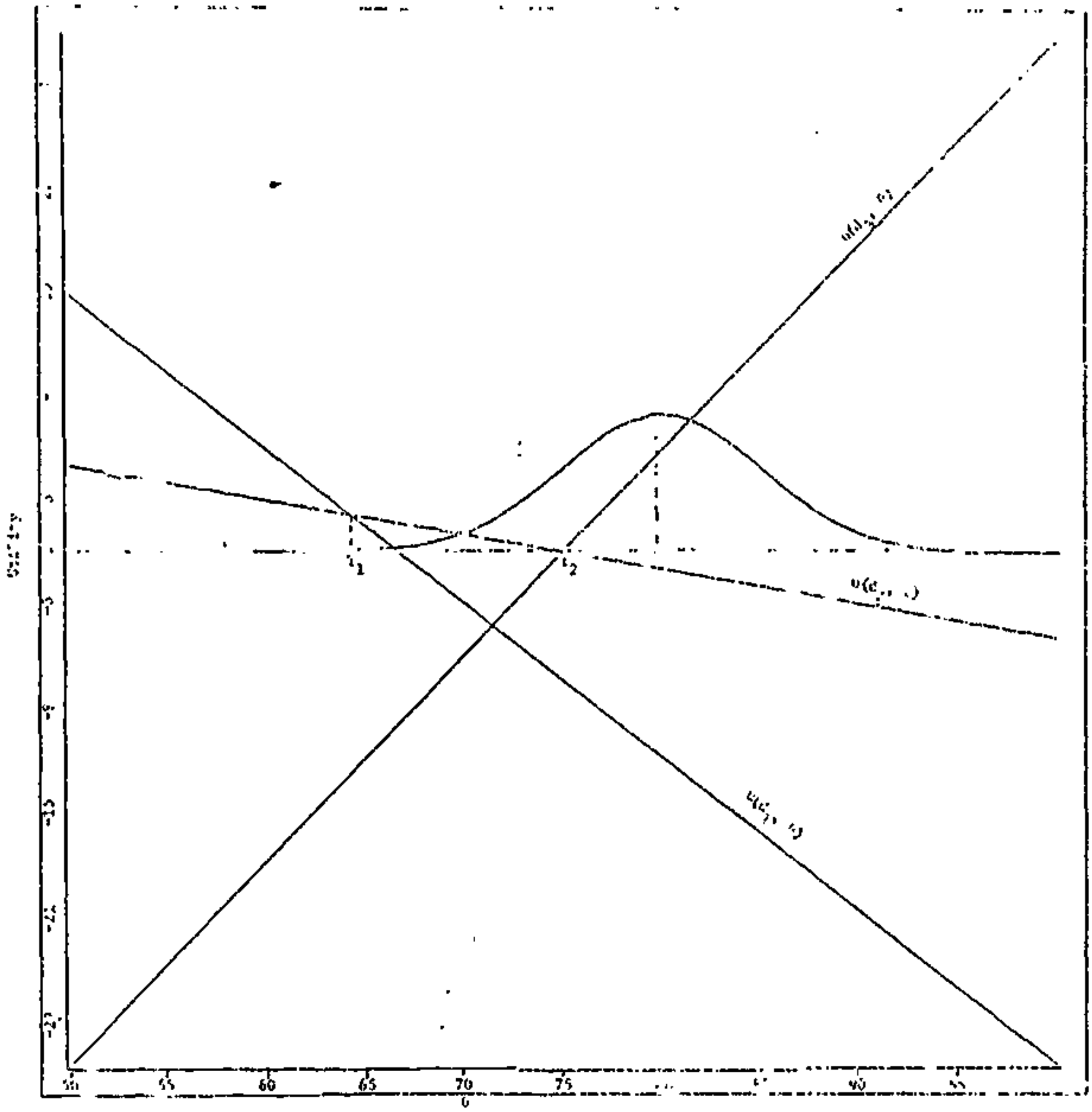
$$e' + f'T_1 = g' + h'T_1$$

and

$$g' + h'T_2 = T_2 \ .$$

Thus, we need only two additional pairs to completely specify the utility function. The resulting linear system of four equations in four unknowns can then be solved for e', f', g', and h'.

When we turn to maximizing expected utility, we now have three equations to consider. In fact, depending upon whether $e' + f'\mu_\theta$, $g' + h'\mu_\theta$, or μ_θ is largest, we choose decision d_1, d_2, or d_3, respectively. Graphically, this is clearly illustrated in Figure 5.1. All the decision maker needs to do is examine the utility of each decision at the mean of the posterior distribution of θ, choosing that decision with the highest value.



<u>Figure 5.1</u>

Linear Utility in a Three-Action Problem

This figure illustrates linear utility in the form of Equation (5.3) with constants e = 60, f = -.9, g = 15, h = -.2, k = -.90, and m = 1.2. As in the two action problem discussed in Section 4.2, the decision depends only upon the utilities of each decision at the posterior mean μ_θ . In this illustration, since $u(d_3, \mu_\theta) > u(d_2, \mu_\theta) > u(d_1, \mu_\theta)$, extensive form analysis will lead our decision maker to choose action d_3. In terms of T_1 and T_2, action d_1 will be taken whenever $\mu_\theta < T_1$; action d_2 will be taken whenever $T_1 < \mu_\theta < T_2$; and, action d_3 will be taken whenever $\mu_\theta > T_2$.

6. Preposterior Analysis

Information is never free. For example, information about the mastery level attained by a particular student is obtained by testing, interviews, or class recitation. Such activities spend the time of the student, spend the time of the teacher, and tie up facilities. If there are only meager rewards and penalties for correct and incorrect decisions, it may be wasteful to purchase information whose cost may exceed the gain in expected utility.

Suppose that a decision maker has (1) a prior distribution on θ, $\rho(\theta)$, (2) a utility function $u(d_i, \theta)$, and (3) a potential experiment which, if carried out, will have outcomes x with model density $p(x|\theta)$. Before collecting the data, the decision maker wants to know the extent to which his efforts are likely to be rewarded. That is, he wants to know whether the additional information contained in the potential experiment is likely to be sufficiently "valuable" to justify obtaining it. Bayesian decision analysis provides the framework of preposterior analysis for studying this question.

The logic of preposterior analysis is simple and can be readily understood by considering the following outline:

(1) The decision maker can attach a "value" to the information contained in his prior, $\rho(\theta)$, by calculating the expected utility of the optimal decision. That is,

$$\text{Value } [\rho(\theta)] = \max_{d_i} \int u(d_i, \theta)\rho(\theta)d\theta \ .$$

(2) Assume for the moment that the experiment has already been carried out and the result x obtained. If this were the case, then analogous to the above, the decision maker could attach a value to the information contained in his posterior. That is,

$$\text{Value } [p(\theta|x)] = \max_{d_i} \int u(d_i, \theta)p(\theta|x)dx \ .$$

(3) Continuing as if the data had already been collected, our decision maker could now calculate the "value added" by the experimental results (i.e., the increase in expected utility after the addition of the data).

$$\text{Value Added} = \max_{d_i} \int u(d_i, \theta)p(\theta|x)d\theta - \max_{d_i} \int u(d_i, \theta)\rho(\theta)d\theta \ .$$

(4) Since, in fact, the experiment has not as yet been executed, of course, the value added cannot be determined. Nevertheless, the decision maker may consider the value added to be a function of the observation random variable X. In the jargon of decision theory, value added considered as a function of the random variable X is referred to as the <u>conditional value of sample (experimental) information</u> and is denoted:

$$v(e, x) = \max_{d_i} \int u(d_i, \theta)p(\theta|x)d\theta - \max_{d_i} \int u(d_i, \theta)\rho(\theta)d\theta \ .$$

It is conditional because it can be calculated only when x is known.

(5) Now since X is a random variable with a probability distribution $p(x)=\int \rho(\theta)p(x|\theta)d\theta$, it is clear that v(e, x) is also a random variable. If the density of v(e, x) were a simple function, it would be useful at this point to examine its location parameters and even credibility intervals. These statistics would describe the decision maker's prior beliefs about the probable increases in utility to be gained from sampling. In most applications, however, the density of v(e, x) is not a simple function. Although this complexity precludes most descriptive indices, in many instances, it will be possible to determine the mean of the distribution of v(e, x). In the decision theory literature, this

mean is commonly referred to as the expected value of sample information and is denoted

$$v(e) = \int v(e, x)p(x)dx$$

where p(x) is the marginal prior density of X.

(6) The decision maker may now compare the expected value of sample information with the "cost" of performing the experiment and judge whether or not the experiment is likely to be worthwhile.

Before illustrating preposterior analysis with a numerical example, one central point must be made. In step (6), the decision maker must compare an expected utility with the cost of obtaining experimental information. It is critical that these two quantities not only be measured in the same units, but also that their respective scales have the same origin. If the expected value of experimental information, v(e), is measured in arbitrary "utile" units while the cost of that information is in dollars and cents, a sensible comparison cannot be expected.

We will illustrate preposterior analysis with an example. We let

$$u(d_i, \theta) = \begin{cases} -(3/5)(\theta - 75) & i = 1 \\ (7/5)(\theta - 75) & i = 2 \end{cases} .$$

As in our previous examples, 75 has been selected as the indifference point between the acts of retaining (d_1) and advancing (d_2) the student. As expected, the advance decision (d_2) is positively related to ability (θ) while the retain decision (d_1) has a negative relationship.

Suppose further that the prior information about θ has been quantified in the form of a normal distribution with mean 78 and variance 36. Recall that in Section 4.2, we demonstrated that with linear utility, the optimum

decision depends only upon whether or not the mean of θ is greater than θ_o . Obviously in this case, with only prior information at the decision maker's disposal, the advance action will result in the highest expected utility. Thus,

$$\begin{aligned} \text{Value}[\rho(\theta)] &= \int (7/5)(\theta - 75)\rho(\theta)d\theta \\ &= u(d_2, \mu_{\theta(\text{prior})}) = (7/5)(\mu_{\theta(\text{prior})} - 75) \end{aligned}$$

where $\mu_{\theta(\text{prior})}$ is the mean of the prior on θ . If the experimenter had carried out the experiment, the highest expected utility using the experimental results to help select the action would be

$$\text{Value }[p(\theta|x)] = \begin{cases} -\int (3/5)(\theta - 75)p(\theta|x)d\theta & \text{if } \mu_{\theta(\text{post})} < 75 \\ \int (7/5)(\theta - 75)p(\theta|x)d\theta & \text{if } \mu_{\theta(\text{post})} \geq 75 . \end{cases}$$

Since the utility function is linear in θ, the expected utility of decision d_i is merely the utility of the expectation or mean of θ . Thus,

$$\text{Value }[p(\theta|x)] = \begin{cases} u(d_1, \mu_{\theta(\text{post})}) = -(3/5)(\mu_{\theta(\text{post})} - 75) \text{ if } \mu_{\theta(\text{post})} < 75 \\ u(d_2, \mu_{\theta(\text{post})}) = (7/5)(\mu_{\theta(\text{post})} - 75) \text{ if } \mu_{\theta(\text{post})} \geq 75 \end{cases}$$

The conditional value of sample information may be given by

$$v(e, x) = \begin{cases} u(d_1, \mu_{\theta(\text{post})}) - u(d_2, \mu_{\theta(\text{prior})}) \text{ if } \mu_{\theta(\text{post})} < 75 \\ u(d_2, \mu_{\theta(\text{post})}) - u(d_2, \mu_{\theta(\text{prior})}) \text{ if } \mu_{\theta(\text{post})} \geq 75 . \end{cases}$$

And the expected value of sample information is given by

$$v(e) = \int\limits_{\substack{\text{all } x \\ \text{such that} \\ \mu_{\theta(post)} < 75}} [u(d_1, \mu_{\theta(post)})]p(x)dx$$

$$+ \int\limits_{\substack{\text{all } x \\ \text{such that} \\ \mu_{\theta(post)} \geq 75}} [u(d_2, \mu_{\theta(post)})]p(x)dx$$

$$- \int u(d_2, \mu_{\theta(prior)})p(x)dx \quad .$$

Thus, in order to evaluate v(e), all our decision maker needs to do is partition the range of x into two subsets: The first containing all x which will lead to a posterior mean which is less than 75, and the second containing those x which force $\mu_{\theta(post)}$ to be greater than 75. If, for the purposes of this example, we assume that the likelihood of our sample may be described by a normal distribution with mean θ and variance 25, then applying Bayes theorem, we see that the posterior distribution of θ is of the form

$$p(\theta|x) \sim N\left[\frac{36\ x + 25*78}{25 + 36}, \frac{36*25}{36 + 25}\right] \quad .$$

And so the relationship $\mu_{\theta(post)} < 75$ is equivalent to the relationship $x < 72.92$. Thus, the expected value of sample information is

$$v(e) = \int_{-\infty}^{72.92} u(d_1, \mu_{\theta(post)})p(x)dx + \int_{72.92}^{+\infty} u(d_2, \mu_{\theta(post)})p(x)dx - u(d_2, \mu_{\theta(prior)}) \int_{-\infty}^{+\infty} p(x)dx \ .$$

Let us pause here for just a moment and examine this equation. The first thing to notice is that we are integrating over a range of test scores X from $-\infty$ to $+\infty$. Conceptually, this may seem a little troublesome, for in most applications, test scores are bounded within a relatively small range.

Recall, however, that we assumed the model density was normal in form. This assumption implies that every x (from $-\infty$ to $+\infty$) has positive probability. Therefore, each x must be considered when taking the expectation. Although admittedly, this is a problem conceptually, in applications it is not very important. As long as the prior distribution on θ is carefully specified, there should be effectively zero probability that x will lie outside its permissible range.

In this particular example, we have $\rho(\theta) \sim N(78, 36)$ and $p(x|\theta) \sim N(\theta, 25)$. Therefore, $p(x) \sim N(78, 61)$. So that in this case, X has very little probability of falling outside the range (55, 101). Returning to our expression for the expected value of sample information, we find two integrals of the form:

$$\int u(d, \mu_{\theta(post)})p(x)dx \; .$$

We know that $u(d_i, \mu_{\theta(post)})$ is linear in $\mu_{\theta(post)}$. Since $\mu_{\theta(post)}$ is linear in x, this implies that $u(d_i, \mu_{\theta(post)})$ is also linear in x . In fact, by substituting $\mu_{\theta(post)} = (36\,x + 25 * 78)/(36 + 25)$, $u(d_i, \mu_{\theta(post)})$ may be written in the linear form

$$u(d_i, \mu_{\theta(post)}) = \begin{cases} -(3/5)(.59x - 43.03) & \text{if } i = 1 \\ \\ (7/5)(.59x - 43.03) & \text{if } i = 2 \; , \end{cases}$$

and so, v(e) can be written as

$$v(e) = \int_{72.92}^{+\infty} (7/5)(.59x-43.03)p(x)dx - \int_{-\infty}^{72.92} (3/5)(.59x-43.03)p(x)dx - u(d_2, \mu_{\theta(prior)})$$

where $p(x) \sim N[78, 61]$. It can be shown that the following relationship holds.

$$\int_{-\infty}^{a} b(cx - d)p(x)dx = \frac{b(c\mu - d)}{100} PR(a) + b\sigma n\left[\frac{\mu - a}{\sigma}; 0, 1\right]$$

where

$$p(x) \sim N[\mu, \sigma^2]$$

$$PR(z) = \text{percentile rank of } z$$

$n\left[\frac{\mu - a}{\sigma}; 0, 1\right]$ = the ordinate or height of the unit normal curve at $\frac{\mu - a}{\sigma}$.

Using this relationship, we may rewrite v(e) once again

$$v(e) = (7/5)\,\frac{(.59\cdot 78 - 43.03)}{100}\,[100 - PR(72.92)] + (7/5)\cdot\sqrt{61}\, n[-.65: \; 0, 1]$$

$$-(3/5)\,\frac{(.59\cdot 78 - 43.03)}{100}\,PR(72.92) - (3/5)\cdot\sqrt{61}\, n[-.65: \; 0, 1]$$

$$-(7/5)(78 - 75)$$

$$= .46 \; .$$

References

Cronbach, L. J., & Gleser, G. C. Psychological Tests and Personnel Decisions. (1st Edition). Urbana: University of Illinois Press, 1957. (2nd Edition, 1965).

De Groot, M. H. Optimal Statistical Decisions. New York: McGraw-Hill, 1970.

Edwards, W., Lindman, H., & Savage, L. J. Bayesian statistical inference for psychological research. Psychological Review, 1963, 70, 193-242.

Hambleton, R., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Research Report No. 54. Iowa City, Iowa: The American College Testing Program, 1972.

Lewis, C., Wang, M., & Novick, M. R. Marginal distributions for the estimation of proportions in m groups. ACT Technical Bulletin No. 13. Iowa City, Iowa: The American College Testing Program, 1973.

Lindley, D. V. Bayesian statistics: A review. CMBS, Regional Conference Series in Applied Mathematics, 1972, No. 2.

Novick, M. R. Computer assisted data analysis. ACT Technical Bulletin No. 3. Iowa City, Iowa: The American College Testing Program, 1972.

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-46.

Novick, M. R., & Jackson, P. H. Statistical Methods for Educational and Psychological Research. New York: McGraw-Hill, 1974.

Raiffa, H., & Schlaifer, R. Applied Statistical Decision Theory. Boston: Division of Research, Harvard Business School, 1961.

Wang, M. Tables of constants for the posterior marginal estimates of proportions in m groups. ACT Technical Bulletin No. 14. Iowa City, Iowa: The American College Testing Program, 1973.

Appendix Number 3

to

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

Appendix Number 3

to

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

The article, "The Estimation of Proportions in m Groups has been removed from this document due to copyright restrictions." The entire article appears in the journal Psychometrika; Vol. 38, No. 1, March 1973.

A C T T E C H N I C A L B U L L E T I N N O. 1 3

MARGINAL DISTRIBUTIONS FOR THE ESTIMATION OF PROPORTIONS IN m GROUPS

by

Charles Lewis

The University of Illinois

Ming-Mei Wang
The American College Testing Program

Melvin R. Novick
The American College Testing Program
and
The University of Iowa

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

May, 1973

1. Introduction

In Individually Prescribed Instruction it has been proposed (Novick, Lewis, and Jackson, 1973; Hambleton and Novick, 1973) that the decision as to whether or not the j-th student has successfully mastered a unit of instruction should be based on the aposteriori probability that his mastery proportion (π_j) is greater than some specified proportion (π_o) and on the losses associated with false-positive and false-negative decisions. It was also proposed that the posterior distribution for each π_j should benefit not only from the prior and sample information on each person j, but also on the collateral information gained from the observations on all other persons.

The rationale for this kind of analysis was first given in an educational context by Kelley (1923, 1927) and later reproposed by Novick (1970), Novick and Jackson (1970), and by Cronbach, Gleser, Nanda, and Rajaratnam (1972). The mathematical structure for the required Bayesian Model II solution was given by Lindley and Smith (1972).

In their recent paper, Novick, Lewis, and Jackson (1973) developed the specific solution for the problem of estimating binomial proportions in m-groups. The observable random variables--proportions of "successes" $p_j = x_j/n_j$, $j = 1, 2, \ldots, m$, where x_j and n_j are respectively the number of successes and the number of observations--were first mapped into a set of new variables g_j by an arc sine transformation. The variables

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U.S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

g_j were then assumed to have a normal density function with mean $\gamma_j = \sin^{-1}\sqrt{\pi_j}$ and variance $v_j = (4n_j + 2)^{-1}$, where γ_j is the corresponding transformed value of the population proportion of "successes" π_j . Then the Bayesian Model II method which is based on the exchangeability theorem of De Finetti (1937) was applied to the analysis of the indirectly observable g_j. In the Individually Prescribed Instruction application, the individual person is treated as the "group" and the test items as the n replications.

The validity of the normality assumption on the distribution of the transformed variables g_j depends on the sample size n_j. If n_j is very small, the normal approximation to the distribution of g_j will not be good. In practice, it was felt that for $n_j \geq 8$ this assumption will be very satisfactory except for the tails. It may also be noted that the domains of the distribution on g_j and γ_j are bounded between 0 and $\frac{\pi}{2}$, while the normal distribution has unbounded domain. We recall that with a uniform prior γ_j, the posterior distribution of γ_j is normal with mean g_j and variance $(4n_j + 2)^{-1}$ under the above appropriate assumptions. Thus, we may wish to check whether the points which are $\pm$ 2 standard deviations from g_j exceed 0 and $\pi/2$, respectively. It was found that for $n_j \geq 6$ and $1 \leq x_j \leq n_j - 1$, the points which are $\pm$ 2 standard deviations from the posterior mean lie within the $(0, \frac{\pi}{2})$ range. This implies that the bounded domain of the distribution of γ_j should not be a major disturbance in considering a normal approximation to its form. We contend that in the m-group procedure, the collateral information provided by other groups would have an equivalent effect of adding more sample observations to the estimation of an individual group's proportion. For this reason, we expect that the violation of normality in the cases of small sample size will not be serious, provided all $n_j \geq 8$. For smaller sample sizes, a logistic transformation introduced by Leonard should be considered, though this will require study.

Regarding the problems of variance stabilization, the Freeman-Tukey (1950) transformation was considered superior to Anscombe's (1948) transformation or the simple arc sine transformation, especially for small sample size n . The condition for the Freeman-Tukey transformation to stabilize the variance is $m\pi(1 - \pi) > 1$. Namely, the true proportion π_j should lie between $\frac{1}{2} - \frac{(1 - 4n^{-1})^{\frac{1}{2}}}{2}$ and $\frac{1}{2} + \frac{(1 - 4n^{-1})^{\frac{1}{2}}}{2}$. In general, this analysis should be very satisfactory provided $m \geq 15$, $n \geq 8$, $\phi_\Gamma \leq .05$, and the above condition is satisfied.

The Bayesian procedure begins with an assumption that the set of transformed values γ_j is a random sample from a normal distribution with mean μ_Γ and variance ϕ_Γ . The analyst's prior beliefa concerning the parameters μ_Γ, ϕ_Γ are partially incorporated into the analysis by specifying prior distributions for them. Specifically, μ_Γ and ϕ_Γ are assumed to be independent, having a uniform and an inverse chi-square (with ν d.f. and parameter λ) density function, respectively. The assumption of a uniform distribution for μ_Γ is more convenient than realistic but does not significantly affect the analysis, provided m is reasonably large. We shall consider this point later.

Under the above distributional assumptions and the Bayesian specifications of one's prior knowledge, the joint probability density function (p.d.f.), $b(\underset{\sim}{g}, \underset{\sim}{\gamma}, \mu_\Gamma, \phi_\Gamma)$, of the vector variables $\underset{\sim}{g}' = (g_1, \ldots, g_j, \ldots, g_m)$, $\underset{\sim}{\gamma}' = (\gamma_1, \ldots, \gamma_j, \ldots, \gamma_m)$ and scalar variables μ_Γ and ϕ_Γ is obtained as:

$$b(\underset{\sim}{g}, \underset{\sim}{\gamma}, \mu_\Gamma, \phi_\Gamma)$$

$$\phi_\Gamma^{-\frac{1}{2}(\nu + m + 2)} \exp\{-\tfrac{1}{2} \sum_i v_i^{-1}(\gamma_i - g_i)^2 + \phi_\Gamma^{-1}(\lambda + \sum_i(\gamma_i - \mu_\Gamma)^2]\} \ . \quad (1.1)$$

Novick, Lewis, and Jackson (1973) arrived at an explicit expression for the posterior density function of $\underset{\sim}{\gamma}$ given $\underset{\sim}{g}$:

$$b(\underset{\sim}{\gamma}|\underset{\sim}{g}) \propto [\lambda + \Sigma(\gamma_i - \gamma.)^2]^{-\frac{1}{2}(\nu + m - 1)} \exp\{-\tfrac{1}{2}[\Sigma v_i^{-1}(\gamma_i - g_i)^2]\} \quad (1.2)'$$

where $\gamma. = \sum_i \gamma_i/m$. Following Lindley's approach, the joint posterior mode $\tilde{\underset{\sim}{\gamma}}$ was taken as the Bayesian modal estimate for $\underset{\sim}{\gamma}$. The components $\tilde{\gamma}_j$ of $\tilde{\underset{\sim}{\gamma}}$ were then used to provide estimates for the group proportions π_j .

The modal estimates of the proportions in m groups taken from the joint posterior distribution from a Bayesian Model II analysis are thought to be more accurate than other estimates obtained from conventional methods. Specifically, the vector estimate of $\underset{\sim}{\gamma}$ should be such as to maximize the probability that <u>all</u> of the components $\tilde{\gamma}_j$ are near the true values γ_j, i.e., the modal estimates minimize zero-one loss in m dimensions. In many applications, however, one's primary concern is to be able to reach certain decisions concerning individual groups (or persons). This would be the case with a component additive-squared error or absolute-error loss function or component threshold loss. Rather than be satisfied with a set of joint estimates, one would, in such situations, like to have marginal means and variances and to make some probability statements about each individual's ability (or a group's level of achievement, etc.). In this context, it is desirable to have knowledge of the marginal distribution of each γ_j . In the present paper, we therefore address ourselves to the problem of describing the posterior marginal distributions of γ_j . To maintain certain mathematical simplicity, the present paper will deal only with the case of equal n . Even with this restriction, the results will still be found applicable in many educational situations (e.g., in assessing students' achievement in a course or instructional unit by administering the same test to each member of a class).

2. Marginal Posterior Distributions for Gammas

An explicit expression for the marginal posterior density function for γ_j does not seem to be obtainable from the joint posterior p.d.f., $b(\underset{\sim}{\gamma}|\underset{\sim}{g})$, of $\underset{\sim}{\gamma}$. However, the joint p.d.f. $b(\underset{\sim}{\gamma}, \underset{\sim}{g}, \mu_\Gamma, \phi_\Gamma)$ given by (1.1), with v_i replaced by $v = (4n + 2)^{-1}$ for equal n, can be integrated with respect to each γ_i ($i \neq j$, $i = 1, \ldots, m$) and μ_Γ to obtain the conditional posterior p.d.f., $b(\gamma_j|\phi_\Gamma, \underset{\sim}{g})$, of γ_j given ϕ_Γ and $\underset{\sim}{g}$:

$$b(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) \propto b(\gamma_j, \phi_\Gamma, \underset{\sim}{g})$$

$$\propto \exp\left[-\frac{m(\phi_\Gamma + v)}{2v(m\phi_\Gamma + v)}\left(\gamma_j - \frac{\phi_\Gamma g_j + vg.}{\phi_\Gamma + v}\right)^2\right], \qquad (2.1)$$

where $g. = \frac{1}{m}\sum_i g_i$. This expression is readily recognized as the kernel of a normal distribution. Thus, the conditional distribution of γ_j given ϕ_Γ and $\underset{\sim}{g}$ is normal with mean

$$\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) = \frac{\phi_\Gamma g_j + vg.}{\phi_\Gamma + v},$$

and variance

$$\mathrm{Var}(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) = \frac{v(\phi_\Gamma + m^{-1}v)}{\phi_\Gamma + v}, \qquad j = 1, \ldots, m.$$

Now if ϕ_Γ can be considered to be known rather precisely, use of the conditional distribution will be justified and requisite constants can be obtained from normal distribution tables. This will occur when m, n are large (e.g., $m > 50$, $n > 30$), as indicated by the computations presented in Table 8 (see section 6). Note that the normal integrations with respect to μ_Γ and the γ_i will be valid, provided the likelihood for these quantities is each near zero outside the admissible range. With respect to μ_Γ, this means that m the number of groups must be large, perhaps $m \geq 15$. In the latter case, this means that the n_j must be moderate, $n_j \geq 8$.

Similarly, we can integrate $b(\underset{\sim}{\gamma}, \underset{\sim}{g}, \mu_\Gamma, \phi_\Gamma)$ w.r.t. $\underset{\sim}{\gamma}$ and μ_Γ to obtain the conditional p.d.f., $b(\phi_\Gamma|\underset{\sim}{g})$, of ϕ_Γ given $\underset{\sim}{g}$ (Hill, 1965; Leonard, 1972):

$$b(\phi_\Gamma|\underset{\sim}{g}) \propto (\phi_\Gamma + v)^{-\frac{m-1}{2}} \cdot \exp\{-\tfrac{1}{2}(\phi_\Gamma + v)^{-1} \sum_{i=1}^{m} (g_i - g.)^2\} \cdot \phi_\Gamma^{-(\frac{\nu}{2}+1)} \exp[-\tfrac{1}{2}\phi_\Gamma^{-1}\lambda] \ . \qquad (2.2)$$

Note that the second factor comes from the prior inverse chi-square distribution of ϕ_Γ, and the first factor is derived from the likelihood of ϕ_Γ given its sufficient statistic $\sum_i (g_i - g.)^2$. This first factor is the kernel of an inverse chi-square density displaced by an amount $-v$. A convenient way to obtain analytically the normalizing constant, mean, and variance for this distribution of $\phi_\Gamma|\underset{\sim}{g}$ does not seem to exist. Hence, direct numerical integration methods will be used for this purpose.

In order to obtain the marginal posterior p.d.f. for γ_j, one would multiply the conditional p.d.f. of γ_j given ϕ_Γ and $\underset{\sim}{g}$ and that of ϕ_Γ given $\underset{\sim}{g}$, as formulated in (2.1) and (2.2), and integrate the result w.r.t. ϕ_Γ . Again, an analytical solution to this problem does not appear to be possible. It is necessary, therefore, to resort to numerical integration methods for computing the marginal posterior means and variances of $\gamma_j|\underset{\sim}{g}$. For this task, the simple form of $b(\gamma_j|\phi_\Gamma, \underset{\sim}{g})$ is helpful in reducing the required computational efforts.

The computational procedure we propose begins with the fact that the r-th moment of $\gamma_j|\underset{\sim}{g}$ equals the expected value (taken over ϕ_Γ given $\underset{\sim}{g}$) of the conditional r-th raw moment of γ_j given ϕ_Γ and $\underset{\sim}{g}$, viz,

$$\mathcal{E}(\gamma_j^r|\underset{\sim}{g}) = \mathcal{E}_{\phi_\Gamma|\underset{\sim}{g}}[\mathcal{E}(\gamma_j^r|\phi_\Gamma, \underset{\sim}{g})] \ . \qquad (2.3)$$

In terms of (2.3), the marginal posterior mean of γ_j is computed by the following equation:

$$\mu_j = \mathcal{E}(\gamma_j|\underset{\sim}{g}) = \mathcal{E}_{\phi_\Gamma}\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) = \mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma g_j + vg.}{\phi_\Gamma + v}\middle|\underset{\sim}{g}\right) = \rho^* g_j + (1 - \rho^*)g. , \tag{2.4}$$

where $\rho^* = \mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma}{\phi_\Gamma + v}\middle|\underset{\sim}{g}\right) = \int_0^\infty\left(\frac{\phi_\Gamma}{\phi_\Gamma + v}\right) b(\phi_\Gamma|\underset{\sim}{g})\, d\phi_\Gamma$, and $1 - \rho^* = \mathcal{E}_{\phi_\Gamma}\left(\frac{v}{\phi_\Gamma + v}\middle|\underset{\sim}{g}\right)$.

We note that $0 < \rho^* < 1$, hence, (2.4) is in fact a weighted average of the values g_j and $g.$. For notational convenience, we shall write ϕ_Γ for $\phi_\Gamma|\underset{\sim}{g}$ in the sequel. Likewise, the expression $\mathcal{E}_{\phi_\Gamma} f(\phi_\Gamma)$ is understood to be the conditional expectation of the function $f(\phi_\Gamma)$ given $\underset{\sim}{g}$.

The posterior variance of γ_j is obtained from the relation:

$$\sigma_j^2 = \mathrm{Var}(\gamma_j|\underset{\sim}{g}) = \mathcal{E}_{\phi_\Gamma}[\mathrm{Var}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})] + \underset{\phi_\Gamma}{\mathrm{Var}}[\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})] . \tag{2.5}$$

Thus, computationally, we use:

$$\begin{aligned}\mathcal{E}_{\phi_\Gamma}[\mathrm{Var}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})] &= \mathcal{E}_{\phi_\Gamma}\left[\frac{v(\phi_\Gamma + m^{-1}v)}{\phi_\Gamma + v}\right] \\ &= v \cdot \mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma}{\phi_\Gamma + v}\right) + m^{-1}v\, \mathcal{E}_{\phi_\Gamma}\left(\frac{v}{\phi_\Gamma + v}\right) \\ &= \rho^* v + (1 - \rho^*)m^{-1}v \quad \text{for all } j = 1, \ldots, m;\end{aligned} \tag{2.6}$$

and

$$\begin{aligned}\underset{\phi_\Gamma}{\mathrm{Var}}[\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})] &= \mathcal{E}_{\phi_\Gamma}\{\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) - \mathcal{E}_{\phi_\Gamma}[\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})]\}^2 \\ &= \mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma g_j + vg.}{\phi_\Gamma + v} - \mu_j\right)^2 \\ &= \mathcal{E}_{\phi_\Gamma}\left[\frac{\phi_\Gamma}{\phi_\Gamma + v}(g_j - \mu_j) + \frac{v}{\phi_\Gamma + v}(g. - \mu_j)\right]^2\end{aligned}$$

$$= (g_j - \mu_j)^2 \mathcal{E}_{\phi_\Gamma}\left[\left(\frac{\phi_\Gamma}{\phi_\Gamma + v}\right)^2\right] + (g. - \mu_j)^2 v^2 \mathcal{E}_{\phi_\Gamma}\left[\frac{1}{(\phi_\Gamma + v)^2}\right]$$

$$+ 2(g_j - \mu_j)(g. - \mu_j)v \, \mathcal{E}_{\phi_\Gamma}\left[\frac{\phi_\Gamma}{(\phi_\Gamma + v)^2}\right] . \tag{2.7}$$

To study the characteristics of the marginal posterior distribution of γ_j, one would also like to compute its coefficient of skewness. For this, we first find the third central moment $Q(\gamma_j|\underset{\sim}{g})$ of γ_j given $\underset{\sim}{g}$ from the general formula:

$$Q(\gamma_j|\underset{\sim}{g}) = \mathcal{E}[\gamma_j - \mathcal{E}(\gamma_j|\underset{\sim}{g})]^3$$

$$= \mathcal{E}_{\phi_\Gamma}[Q(\gamma_j|\phi_\Gamma, \underset{\sim}{g})] + Q_{\phi_\Gamma}[\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})]$$

$$+ 3 \underset{\phi_\Gamma}{\mathrm{Cov}}[\mathrm{Var}(\gamma_j|\phi_\Gamma, \underset{\sim}{g}), \mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})] , \tag{2.8}$$

where Cov denotes a covariance. In the present case, $Q(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) = 0$ since the conditional posterior distribution of γ_j given ϕ_Γ and $\underset{\sim}{g}$ is normal [see equation (2.1)]. Furthermore,

$$Q_{\phi_\Gamma}[\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})] = \mathcal{E}_{\phi_\Gamma}[\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) - \mathcal{E}_{\phi_\Gamma}\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})]^3$$

$$= \mathcal{E}_{\phi_\Gamma}\left[\frac{\phi_\Gamma g_j + v g.}{\phi_\Gamma + v} - \mu_j\right]^3$$

$$= \sum_{\ell=0}^{3} \binom{3}{\ell} (g_j - \mu_j)^\ell [v(g. - \mu_j)]^{3-\ell} \cdot \mathcal{E}_{\phi_\Gamma}\left[\frac{\phi_\Gamma^\ell}{(\phi_\Gamma + v)^3}\right] , \tag{2.9}$$

where $Q_{\phi_\Gamma}[\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})]$ is the third central moment (w.r.t. ϕ_Γ) of the conditional expectation $\mathcal{E}(\gamma_j|\phi_\Gamma, \underset{\sim}{g})$, and

$$\underset{\phi_\Gamma}{\mathrm{Cov}}[\mathrm{Var}(\gamma_j|\phi_\Gamma, g), \mathcal{E}(\gamma_j|\phi_\Gamma, g)]$$

$$= \mathcal{E}_{\phi_\Gamma}\left[\frac{v(\phi_\Gamma + v/m)}{\phi_\Gamma + v}\left(\frac{\phi_\Gamma g_j + vg.}{\phi_\Gamma + v} - \mu_j\right)\right]$$

$$= v(g_j-\mu_j)\cdot\mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma}{\phi_\Gamma+v}\right)^2 + \left[m^{-1}v^2(g_j-\mu_j)+v^2(g.-\mu_j)\right]\mathcal{E}_{\phi_\Gamma}\left[\frac{\phi_\Gamma}{(\phi_\Gamma+v)^2}\right]$$

$$+ m^{-1}v^3(g. - \mu_j)\cdot\mathcal{E}_{\phi_\Gamma}\left[\frac{1}{(\phi_\Gamma + v)^2}\right] . \qquad (2.10)$$

Hence, in terms of equations (2.5) to (2.10), one finds for the coefficient of skewness δ_j of the marginal posterior distribution of γ_j given g:

$$\delta_j = Q(\gamma_j|g)/[\mathrm{Var}(\gamma_j|g)]^{3/2} .$$

In summary, it is seen from equations (2.4) to (2.10) that given the expectations with respect to ϕ_Γ of the functions $\phi_\Gamma^\ell(\phi_\Gamma + v)^{-k}$ $(0 \leq \ell \leq k,\ k = 1, 2, 3)$ of ϕ_Γ and the indirectly observable vector g, the descriptive statistics of our interest--the mean μ_j, variance σ_j^2, and index of skewness δ_j--for the marginal posterior distributions of the γ_j can be easily computed. To obtain the values $\mathcal{E}_{\phi_\Gamma}[\phi_\Gamma^\ell(\phi_\Gamma + v)^{-k}]$, we use numerical integration methods. First, the right-hand side (r.h.s.) of (2.2) is integrated w.r.t. ϕ_Γ $(0 < \phi_\Gamma < \infty)$, and the reciprocal of the resulting value is taken to give the proportionality constant for $b(\phi_\Gamma|g)$ in (2.2). The particular integration algorithm adopted here is one which applies Simpson's rule and uses local parabolic fitting to the curve being integrated in computing the partitioned integral over a small range of the argument (in this case, ϕ_Γ). For detailed information, the reader may refer to Ralston (1965, p. 119).

The next step involves computing the expectations of $f_{\ell,k}(\phi_\Gamma) = \phi_\Gamma^\ell(\phi_\Gamma + v)^{-k}$. The same integration algorithm described above is employed to obtain $\mathcal{E}_{\phi_\Gamma} f_{\ell,k}(\phi_\Gamma) = \int_0^\infty f_{\ell,k}(\phi_\Gamma) b(\phi_\Gamma|\underset{\sim}{g}) d\phi_\Gamma$. The mean, variance, and index of skewness for the marginal posterior distribution of γ_j are then obtained via equations (2.4) to (2.10). A Fortran IV program, MARPRO, was written to carry out all these computations.

Finally, the exact posterior probability $\text{prob}(\pi_j \geq \pi_o|\underset{\sim}{g})$ that the j-th group's proportion is greater than or equal to some prespecified cutting point π_o given the observed vector $\underset{\sim}{g}$ can also be calculated. Explicitly,

$$
\begin{aligned}
&\text{prob}(\pi_j \geq \pi_o|\underset{\sim}{g}) \\
&= \text{prob}(\gamma_j > \gamma_o|\underset{\sim}{g}) \\
&= \int_{\gamma_o}^\infty b(\gamma_j|\underset{\sim}{g}) d\gamma_j \\
&= \int_{\gamma_o}^\infty \int_0^\infty b(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) b(\phi_\Gamma|\underset{\sim}{g}) d\phi_\Gamma d\gamma_j \\
&= \int_0^\infty \left[\int_{\gamma_o}^\infty b(\gamma_j|\phi_\Gamma, \underset{\sim}{g}) d\gamma_j \right] \cdot b(\phi_\Gamma|\underset{\sim}{g}) d\phi_\Gamma \quad ,
\end{aligned}
$$

where $\gamma_o = \sin^{-1}(\sqrt{\pi_o})$ is the arc sine transformation of π_o . The inner integral for given ϕ_Γ is recalled to be the upper end cumulative normal probability since $b(\gamma_j|\phi_\Gamma, \underset{\sim}{g})$ is a normal density. The outer integral (w.r.t. ϕ_Γ) is obtained using the same numerical integration algorithm described earlier in this section. The program MARPRO also provides these probabilities with various values of π_o (for $.95 \geq \pi_o \geq .05$ in steps of .05, terminating with a value π_o for which $\text{prob}(\pi_j \geq \pi_o) \geq .99$).

3. Marginal Mean Estimates as Compared to Joint Modal Estimates

In the Novick, Lewis, and Jackson (1973) paper, the joint posterior modal estimate ($\tilde{\gamma}_j$) for γ_j (the arc sine transformation of the proportion π_j) is obtained as a weighted average of g_j (the corresponding transformation of the observed proportion p_j) and the average $\tilde{\gamma}.$ of the estimated values $\tilde{\gamma}_j$ in m-groups. Explicitly,

$$\tilde{\gamma}_j = \rho_j g_j + (1 - \rho_j)\tilde{\gamma}. \, , \tag{3.1}$$

where

$$\rho_j = \left[\frac{\lambda + \sum_i(\tilde{\gamma}_i - \tilde{\gamma}.)^2}{m + \nu - 1}\right] \Bigg/ \left[\frac{\lambda + \sum_i(\tilde{\gamma}_i - \tilde{\gamma}.)^2}{m + \nu - 1} + v_j\right] \text{ and } \tilde{\gamma}. = m^{-1} \sum_i \tilde{\gamma}_i \, .$$

In the case where all m groups have same sample sizes, $n_i = n$, equation (3.1) can be simplified as

$$\tilde{\gamma}_j = \rho g_j + (1 - \rho)\tilde{\gamma}. = \rho g_j + (1 - \rho)g. \, , \tag{3.2}$$

since now $\tilde{\gamma}. = g.$. Here ρ can be obtained as the solution of a cubic equation [Novick, Lewis, and Jackson, 1973, p. 37, (6.18)]. It may be recalled that a parallel expression for the marginal posterior mean (μ_j) of γ_j [equation (2.4)] was obtained in the previous section. There the weight ρ^* is the conditional mean (w.r.t. ϕ_Γ) of $\frac{\phi_\Gamma}{\phi_\Gamma + v}$ given $\underset{\sim}{g}$. (All estimates concerning us hereafter are understood to be the posterior estimates so that the word "posterior" will be omitted in the sequel.)

Returning to (3.2), we may write

$$\rho = \frac{\tilde{\phi}_\Gamma}{\tilde{\phi}_\Gamma + v} \quad ,$$

where $\tilde{\phi}_\Gamma = [\lambda + \sum_i(\tilde{\gamma}_i - \tilde{\gamma}.)^2]/(m + \nu - 1)$ is an estimate of ϕ_Γ . Thus, both equations (3.2) and (2.4) are special forms of the Kelley type formula (Kelley, 1927). The only difference is that ρ is an estimate of the reliability $R[= \phi_\Gamma(\phi_\Gamma + v)^{-1}]$ based on an estimate of the variance ϕ_Γ, while ρ^* is the expected value (over ϕ_Γ) of R given g (i.e., a Bayesian mean estimate of R w.r.t. ϕ_Γ).

At this point, we are interested in comparing the marginal mean estimates $\tilde{\tilde{\gamma}}_j = \mu_j$ of γ_j (or equivalently $\tilde{\tilde{\pi}}_j$ of the proportions π_j) with their joint modal estimates ($\tilde{\gamma}_j$, or, equivalently, $\tilde{\pi}_j$). This comparison relies solely on the relative magnitudes of ρ and ρ^* . We have found rom our numerical investigation that ρ^* is substantially larger than ρ for moderate n . This means that the marginal mean estimates are less regressed towards the common value g. than the joint modal estimates. Similarly, we would expect that the marginal modal estimates would be less regressed to the common g. than the joint modal estimates. In particular, the marginal modal estimates coincide with the marginal mean estimates when the marginal distributions are unimodal and symmetric. In the present context, the marginal distribution of γ_j given g is unimodal and nearly symmetric. More discussions on the shape of these distributions will be given in section 5. To elaborate the above results, let us rewrite equations (2.4) and (3.2) as:

$$\mathcal{E}(\gamma_j|g) = \tilde{\tilde{\gamma}}_j = g. + \rho^*(g_j - g.) \ ,$$

and

$$\tilde{\gamma}_j = g. + \rho(g_j - g.) \quad .$$

Then it is obvious that if for a particular group j its observed g_j is greater than g., we have $\tilde{\tilde{\gamma}}_j > \tilde{\gamma}_j$, conversely for $g_j \leq g.$ we find $\tilde{\tilde{\gamma}}_j \leq \tilde{\gamma}_j$. In terms of proportions, we obtain

$$\tilde{\tilde{\pi}}_j > \tilde{\pi}_j \qquad \text{if } g_j > g.$$

and

$$\tilde{\tilde{\pi}}_j < \tilde{\pi}_j \qquad \text{if } g_j < g. \,, \tag{3.3}$$

where $\tilde{\tilde{\pi}}_j$ and $\tilde{\pi}_j$ are estimates of the proportion π_j based on $\tilde{\tilde{\gamma}}_j$ and $\tilde{\gamma}_j$, respectively. Numerical illustrations of the relation (3.3) are given in section 6.

The reader is again reminded that the problem of estimation is closely linked with the concept of loss function. Different estimators are chosen for different loss functions. The substantial discrepancies found between the joint modal estimates and the marginal mean estimates of γ_j suggest that the defined loss function and the kind of decision (an overall decision for all groups or decisions to be made on individual groups separately) are important in the present estimation problem. If one is primarily interested in making an overall decision for all groups (persons, in many applications to the educational assessment practices) and zero-one loss is chosen, he would take the joint modal estimates. On the other hand, if individual decisions are the main concern and squared-error loss is considered appropriate, he would choose to use the marginal mean estimates. For individual decisions with zero-one loss the marginal modes would be the ideal estimators. However, in the present context, these marginal modes would likely be close to the marginal means.

One final comment on the effect of sample size n . The reliability $R = \phi_\Gamma(\phi_\Gamma + v)^{-1}$ increases as n becomes larger, since $v = (4n + 2)^{-1}$ decreases. Hence, both ρ and ρ^* (being estimators of R) are also expected to increase with n . In the limit $(n \to \infty)$, both will approach unity. That is, our estimates will be based completely on the observed values g_j . For this same reason, the estimate $\tilde{\tilde{\gamma}}_j$ and $\tilde{\gamma}_j$ (or, $\tilde{\tilde{\pi}}_j$ and π_j) will differ less for larger n . On the other hand, as m increases, more collateral information is available. One would then be likely to shift more weight to the common value in obtaining estimates for γ_j . Detailed numerical examples are provided in section 6.

4. Some Limiting Distributions for Gamma

In the Bayesian estimation of m-group proportions, it is said that the remaining m - 1 groups in effect provide some sort of "prior information" (strictly speaking, collateral information since it is not obtained prior to analysis) for estimating the proportions in an individual group. In view of this statement one may hope to find an approximate expression for $b(\gamma_j|\underset{\sim}{g})$ by first working with the posterior p.d.f. $b(\gamma_j|\underset{\sim}{\gamma}^*, \underset{\sim}{g})$ of γ_j given $\underset{\sim}{g}$, assuming the vector $\underset{\sim}{\gamma}^{*\prime} = (\gamma_1, \gamma_2, \ldots, \gamma_{j-1}, \gamma_{j+1}, \ldots, \gamma_m)$ is known. This posterior distribution can be obtained from equation (1.2) by making the substitutions:

$$\gamma_j - \gamma_{\cdot} = \frac{m-1}{m}(\gamma_j - \gamma_{\cdot}^*)$$

and

$$\sum_i(\gamma_i - \gamma_{\cdot})^2 = \sum_{i\neq j}(\gamma_i - \gamma_{\cdot}^*)^2 + \frac{m-1}{m}(\gamma_j - \gamma_{\cdot}^*)^2 ,$$

where

$$\gamma_{\cdot}^* = (m-1)^{-1}\sum_{i\neq j}\gamma_i .$$

Thus, we arrive at

$$b(\gamma_j|\underset{\sim}{\gamma}^*, \underset{\sim}{g}) \propto [\lambda + \sum_{i\neq j}(\gamma_i - \gamma_{\cdot}^*)^2 + \frac{m-1}{m}(\gamma_j - \gamma_{\cdot}^*)^2]^{-\frac{1}{2}(\nu + m - 1)}$$

$$\cdot\{\exp[-\tfrac{1}{2}v_j^{-1}(\gamma_j - g_j)^2]\} \qquad (4.1)$$

since

$$b(\gamma_j|\underset{\sim}{\gamma}^*, \underset{\sim}{g}) \propto b(\underset{\sim}{\gamma}|\underset{\sim}{g}) .$$

The second factor of the r.h.s. of (4.1) is the likelihood of γ_j given g_j and the first factor can be regarded as the contribution from the prior information about γ_j provided by $\underset{\sim}{\gamma}^*$ (in addition to λ and ν, of course).

Therefore, the posterior distribution of γ_j given g is explicitly available if γ^* is indeed known. In reality, it is unlikely to know $\underset{\sim}{\gamma}^*$ beforehand. At first thought, one may be tempted to substitute the joint modal estimates $\tilde{\gamma}_i$ (which are relatively easy to obtain as given in Novick, Lewis, and Jackson, 1973) for γ_i $(i \neq j)$ in (4.1) to find an approximation for $b(\gamma_j|g)$. This expedient step is appealing because only the mean $\gamma^*_{\cdot}$ and sum of squared deviations $\sum_{i\neq j} (\gamma_i - \gamma^*_{\cdot})^2$ enter to the density function (4.1). This approach was tried but found to be insufficiently precise.

Returning to equation (4.1), it is noted that the first term of its r.h.s. is the kernel of a nonstandardized t-distribution with d.f. $\nu^* = \nu + m - 2$ and parameters $\zeta = \gamma^*_{\cdot}$, $\kappa = m(m - 1)^{-1}[\lambda + \sum_{i\neq j} (\gamma_i - \gamma^*_{\cdot})^2]$. (See Novick and Jackson, 1974.) When $m \to \infty$, this t-distribution approaches a normal distribution with mean $\gamma^*_{\cdot}$ and variance $\phi^* = \{\lambda + \sum_{i\neq j} (\gamma_i - \gamma^*_{\cdot})^2\}/(\nu + m - 4)$. Consequently, $b(\gamma_j|\underset{\sim}{\gamma}^*, g, m \to \infty)$, being proportional to the product of a normal likelihood and a normal prior density, is itself a normal density. We conclude from this standard Bayesian result that the limiting $(m \to \infty)$ posterior distribution of γ_j given g, for known $\underset{\sim}{\gamma}^*$, is a normal distribution with mean $\mathcal{E}(\gamma_j|\underset{\sim}{\gamma}^*, g, m \to \infty) = (\phi^* g_j + v_j \gamma^*_{\cdot})/(\phi^* + v_j)$ and variance $\mathrm{Var}(\gamma_j|\underset{\sim}{\gamma}^*, g, m \to \infty) = v_j \phi^*/(\phi^* + v_j)$. Unfortunately, this limiting distribution is not very useful in practice since $\underset{\sim}{\gamma}^*$ is not typicallly known.

A second related limiting distribution which might be of interest is that of γ_j given g when both m and n tend to infinity. For equal sample size n, integrating the joint p.d.f. $b(\underset{\sim}{\gamma}, g, \mu_\Gamma, \phi_\Gamma)$ [equation (1.1)] w.r.t. each γ_i $(i \neq j)$ and then w.r.t. μ_Γ, we obtain the joint posterior p.d.f. of γ_j and ϕ_Γ given g:

$$b(\gamma_j, \phi_\Gamma | g) \propto \exp[-\tfrac{1}{2}v^{-1}(\gamma_j - g_j)^2]\ f_1(\phi_\Gamma, \gamma_j)\ f_2(\phi_\Gamma)\ f_3(\phi_\Gamma) \ , \qquad (4.2)$$

where

$$f_1(\phi_\Gamma, \gamma_j) = (\phi_\Gamma + \frac{v}{m})^{-\frac{1}{2}} \exp\{-\tfrac{1}{2} \frac{m-1}{m} (\phi_\Gamma + \frac{v}{m})^{-1}(\gamma_j - g^*_\cdot)^2\} \ ,$$

$$f_2(\phi_\Gamma) = (\phi_\Gamma + v)^{-\frac{m-2}{2}} \exp\{-\tfrac{1}{2}(\phi_\Gamma + v)^{-1} \sum_{i \neq j} (g_i - g^*_\cdot)^2\} \ ,$$

and

$$f_3(\phi_\Gamma) = b(\phi_\Gamma) = \phi_\Gamma^{-\frac{1}{2}(\nu + 2)} \exp(-\tfrac{1}{2}\lambda\phi_\Gamma^{-1}) \ .$$

(Note: $g^*_\cdot = (m - 1)^{-1} \sum_{i \neq j} g_i$ and $b(\phi_\Gamma)$ is the prior p.d.f. of ϕ_Γ).

As $n \to \infty$, the contribution of $v(\to 0)$ and $m^{-1}v(\to 0)$ to f_1 and f_2 in (4.2), relative to that of ϕ_Γ will become negligible. It follows that $f_1(\phi_\Gamma, \gamma_j)f_2(\phi_\Gamma)f_3(\phi_\Gamma)$ may be approximated by

$$\phi_\Gamma^{-\frac{1}{2}(\nu + m + 1)} \exp\{-\tfrac{1}{2}\phi_\Gamma^{-1}[\lambda + \sum_{i \neq j} (g_i - g^*_\cdot)^2 + \frac{m-1}{m} (\gamma_j - g^*_\cdot)^2]\} \ ,$$

so that

$$\int_0^\infty f_1(\phi_\Gamma, \gamma_j)\cdot f_2(\phi_\Gamma)\cdot f_3(\phi_\Gamma)\,d\phi_\Gamma$$

$$\propto [\lambda + \sum_{i \neq j} (g_i - g^*_\cdot)^2 + \frac{m-1}{m} (\gamma_j - g^*_\cdot)^2]^{-\frac{1}{2}(\nu + m - 1)} ,$$

as $n \to \infty$. Thus, $b(\gamma_j | g) = \int_0^\infty b(\gamma_j, \phi_\Gamma | g)\,d\phi_\Gamma$ is approximately proportional to

$$\exp[-\tfrac{1}{2}v^{-1}(\gamma_j - g_j)^2]\ [\lambda + \sum_{i \neq j} (g_i - g^*_\cdot)^2 + \frac{m-1}{m} (\gamma_j - g^*_\cdot)^2]^{-\frac{1}{2}(\nu + m - 1)} \ , \qquad (4.3)$$

when $n \to \infty$.

If we further let $m \to \infty$, the second factor of the expression (4.3), being the kernel of an unstandardized t density, approaches a normal density

with mean $g_{\cdot}^{*}$ and variance $\hat{\phi} = \{\lambda + \sum_{i \neq j} (g_i - g_{\cdot}^{*})^2\}/(\nu + m - 4)$. The expression (4.3) is parallel to the r.h.s. of (4.1) with $\gamma_i (i \neq j)$ and $\gamma_{\cdot}^{*}$ replaced by g_i $(i \neq j)$ and $g_{\cdot}^{*}$. On the same ground discussed in connection with $b(\gamma_j | \underset{\sim}{\gamma}^{*}, \underset{\sim}{g}, m \to \infty)$, it is then obvious that, when $m, n \to \infty$, the posterior marginal distribution of γ_j given $\underset{\sim}{g}$ is normal with mean $\mathcal{E}(\gamma_j | g, m, n \to \infty) = (\hat{\phi} g_j + v g_{\cdot}^{*})(\hat{\phi} + v)^{-1}$ and variance $\mathrm{Var}(\gamma_j | \underset{\sim}{g}, m, n \to \infty) = \hat{\phi} v (\hat{\phi} + v)^{-1}$. This simple form of the limiting distribution for $b(\gamma_j | \underset{\sim}{g})$ suggests exploring the possibility of a normal approximation to the exact posterior marginal distribution of γ_j (see section 5).

In passing, we note that another attempt to approximate $b(\gamma_j | \underset{\sim}{g})$ by substituting the kernel of inverse chi-square densities $f_1^{*}(\phi_\Gamma)$ and $f_2^{*}(\phi_\Gamma)$ of ϕ_Γ, having modes same as the modes of ϕ_Γ in f_1 and f_2, for $f_1(\phi_\Gamma, \gamma_j)$ and $f_2(\phi_\Gamma)$ in (4.2) was also made. In this case,

$$f_1^{*}(\phi_\Gamma) = \phi_\Gamma^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2} \phi_\Gamma^{-1} \left[\frac{m-1}{m} (\gamma_j - g_{\cdot}^{*})^2 - \frac{v}{m}\right]\right\},$$

and

$$f_2^{*}(\phi_\Gamma) = \phi_\Gamma^{-\frac{m-2}{2}} \exp\{-\tfrac{1}{2} \phi_\Gamma^{-1} [\sum_{i \neq j} (g_i - g_{\cdot}^{*})^2 - (m-2)v]\} .$$

Thus, we have

$$\int_0^\infty [f_1^{*}(\phi_\Gamma)] \cdot [f_2^{*}(\phi_\Gamma)] \cdot [f_3(\phi_\Gamma)] \, d\phi_\Gamma$$

$$\propto \{\lambda + [\sum_{i \neq j} (g_i - g_{\cdot}^{*})^2 - (m-2)v] + \frac{m-1}{m} (\gamma_j - g_{\cdot}^{*})^2 - \frac{v}{m}\}^{-\frac{\nu + m - 1}{2}} . \tag{4.4}$$

The result of replacing f_1^{*} and f_2^{*} for f_1 and f_2 in $b(\gamma_j | \underset{\sim}{g})$ is then

$$b^{*}(\gamma_j | \underset{\sim}{g}) = \exp[-\tfrac{1}{2} v^{-1} (\gamma_j - g_j)^2]$$

$$\cdot \{\lambda + [\sum_{i \neq j} (g_i - g_{\cdot}^{*})^2 - (m-2)v] + \frac{m-1}{m} (\gamma_j - g_{\cdot}^{*})^2 - \frac{v}{m}\}^{-\frac{1}{2}(\nu + m - 1)} . \tag{4.5}$$

Examining (4.5), we find that $b^*(\gamma_j|\underset{\sim}{g})$ is of the same form of $b(\gamma_j|\underset{\sim}{\gamma}^*, \underset{\sim}{g})$ in (4.1) except an extra term $\frac{v}{m}$ and that $\sum_{i\neq j} (\gamma_i - \gamma_\cdot^*)^2$ is replaced by the quantity $\sum_{i\neq j} (g_i - g_\cdot^*)^2 - (m - 2)v$. Again, as m, $n \to \infty$, $b^*(\gamma_j|\underset{\sim}{g})$ approaches a normal density.

The approximation of $b(\gamma_j|\underset{\sim}{g})$ by $b^*(\gamma_j|\underset{\sim}{g})$ is in effect a special case of those by $b(\gamma_j|\underset{\sim}{\gamma}^*, \underset{\sim}{g})$. This is so because if we adopt Jackson's (1972) proposed estimates of $\sum_i(g_i - g_\cdot)^2 - (m - 1)v$ for $\sum_i(\gamma_i - \gamma_\cdot)^2$ and $g_\cdot$ for $\gamma_\cdot$ (in the present context, $\sum_{i\neq j} (\gamma_i - \gamma_\cdot^*)^2$ is estimated by $\sum_{i\neq j} (g_i - g_\cdot^*)^2 - (m - 2)v$ and $\gamma_\cdot^*$ by $g_\cdot^*$) and ignore the term v/m (which should be negligible even for moderate m and n), we can treat $b^*(\gamma_j|\underset{\sim}{g})$ as derived from $b(\gamma_j|\underset{\sim}{\gamma}^*, \underset{\sim}{g})$. Though seemingly appealing, this effort to obtain an approximation for $b(\gamma_j|\underset{\sim}{g})$ also fails. However, it is a comfort to learn that a normal approximation to $b(\gamma_j|\underset{\sim}{g})$ has been found satisfactory. This approximation is discussed in the next section.

5. A Normal Approximation to the Posterior Marginal Distribution of Gamma

Searching for an approximation to the posterior marginal distribution of γ_j, we carefully studied the shape of its exact form. The index of skewness shows that it is only slightly skewed and that the skewness can therefore be ignored. In general, the marginal distribution of γ_j given $\underset{\sim}{g}$ is positively skewed if the observed g_j is greater than $g.$, and negatively skewed if $g_j < g.$. The exact ordinates $b(\gamma_j|\underset{\sim}{g})$ of its density curve at those points within the range of $\pm$ 2 standard deviations (σ_j) from its mean (μ_j) in steps of .5 σ_j were evaluated by numerical integrations. The results invariably indicate a pattern of unimodality. The density is higher in the central region around the mean and decreases as γ_j moves away from μ_j . This suggests a good possibility of approximating this density curve by a normal curve except perhaps in the tails. Also, we recall that indeed it has a normal density as its limiting form. We, therefore, compared these exact ordinates $b(\gamma_j|\underset{\sim}{g})$ with the corresponding ordinates of a normal curve whose mean and standard deviation coincide with μ_j and σ_j of the exact distribution for γ_j given $\underset{\sim}{g}$. These comparisons did bear out our conjecture that the normal approximation is a promising approach.

As an example, a data set which was the result of a test of 12 items administered to 35 children was used. There were 3, 4, 5, 12, and 11 persons with 8, 9, 10, 11, and 12 correct answers, respectively. Columns 2 and 3 of Table 1 contain the ordinates $b(\gamma_j|\underset{\sim}{g})$ of the exact distribution of $\gamma_j|\underset{\sim}{g}$ for the persons j having 10 correct answers, and those of the corresponding normal curve. The points $\gamma_c = \gamma_j + c\sigma_j$, where c takes the values from -2.5 to +2.5 in steps of .5, were included. We remark that in making decisions, the relevant information is often based on the cumulative probability rather than the density itself.

For this reason, columns 4 and 5 present the exact cumulative probabilities $\mathrm{Prob}(\gamma_j \leq \gamma_c | \underset{\sim}{g})$ and the corresponding normal approximations. It is seen that the discrepancies between the exact and approximate figures are less than .01. In most practical applications, this accuracy should prove to be entirely satisfactory.

The currently available program MARPRO provides the exact probability $\mathrm{Prob}(\pi_j \geq \pi_o | \underset{\sim}{g}) = \mathrm{Prob}(\gamma_j \geq \gamma_o | \underset{\sim}{g})$ as well as its normal approximation. This normal approximation has been found to be very adequate. The differences are, in fact, less than .005 in nearly all cases. More numerical illustrations are given in the next section.

Frequently, one is interested in finding the 100α percentage points for gamma. They are difficult to evaluate directly from the exact distribution of $\gamma_j | \underset{\sim}{g}$. However, one could find the approximate 100α percentage point $\gamma_{\alpha j}$ for γ_j using the unit normal curve, since the normal approximation is usually expected to be sufficiently accurate. For this purpose, we now derive an expression for $\sigma_j^2 \left[= \mathrm{Var}(\gamma_j | \underset{\sim}{g})\right]$ in terms of $\rho^* \left[= \mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma}{\phi_\Gamma + v} | \underset{\sim}{g}\right)\right]$ and

$$\sigma^{*2} = \underset{\phi_\Gamma}{\mathrm{Var}} \left(\frac{\phi_\Gamma}{\phi_\Gamma + v} | \underset{\sim}{g}\right) . \tag{5.1}$$

First, we find

$$\begin{aligned} &\underset{\phi_\Gamma}{\mathrm{Var}} \left[\mathcal{E}(\gamma_j | \phi_\Gamma, \underset{\sim}{g})\right] \\ &= \underset{\phi_\Gamma}{\mathrm{Var}} \left(\frac{\phi_\Gamma g_j + v g.}{\phi_\Gamma + v} \right) \\ &= \underset{\phi_\Gamma}{\mathrm{Var}} \left[g. + \frac{\phi_\Gamma}{\phi_\Gamma + v} (g_j - g.)\right] \end{aligned}$$

$$= (g_j - g.)^2 [\operatorname*{Var}_{\phi_\Gamma} (\frac{\phi_\Gamma}{\phi_\Gamma + v} | \underset{\sim}{g})]$$

$$= (g_j - g.)^2 \sigma^{*2} . \qquad (5.2)$$

Applying equations (2.5), (2.6), and (5.1), we then have

$$\sigma_j^2 = \operatorname{Var}(\gamma_j | \underset{\sim}{g}) = v \left[\frac{(m - 1)\rho^* + 1}{m} \right] + (g_j - g.)^2 \sigma^{*2} . \qquad (5.3)$$

The advantage of using the formula (5.3) is that in order to find the approximate 100α percentage points of γ_j for all groups, we need only to compute ρ^* and σ^{*2} . Given $b(\phi_\Gamma|\underset{\sim}{g})$ in equation (2.2), ρ^* and σ^{*2} are easy to compute by numerical integrations, for known m, n, λ, ν, and $\sum_i (g_i - g.)^2$.

Now, if one has ρ^* and σ^{*2} available, the 100α percentage point $\gamma_{\alpha j}$ for $\gamma_j|\underset{\sim}{g}$ can be obtained with the help of a standard normal table. Thus, let

$$\operatorname{prob}(z \leq z_\alpha) = \alpha ,$$

where z is a standard normal variate, one finds

$$\gamma_{\alpha j} = z_\alpha \left[\frac{1 + (m - 1)\rho^*}{(4n + 2)m} + (g_j - g.)^2 \sigma^{*2} \right]^{\frac{1}{2}} + \rho^*(g_j - g.) + g.$$

such that

$$\operatorname{prob}(\gamma_j \leq \gamma_{\alpha j} | \underset{\sim}{g}) = \alpha .$$

The sine-squared transformation of $\gamma_{\alpha j}$ can then be taken as the approximate 100α percentage point for π_j, viz., $\pi_{\alpha j} = \sin^2 \gamma_{\alpha j}$ for which $\operatorname{prob}(\pi_j \leq \pi_{\alpha j} | \underset{\sim}{g}) = \alpha$. Similarly, knowing ρ^* and σ^{*2}, one could evaluate the approximate probability of $\gamma_j \leq \gamma_o$ given $\underset{\sim}{g}$ from the normal table:

$$\mathrm{Prob}(\gamma_j \le \gamma_o | \underset{\sim}{g}) \approx \mathrm{Prob}(z \le z_o)$$

where

$$z_o = \frac{\gamma_o - \rho^*(g_j - g.) - g.}{\left[\frac{1 + (m - 1)\rho^*}{(4n + 2)m} + (g_j - g.)^2 \sigma^{*2}\right]^{\frac{1}{2}}} .$$

Before leaving this discussion, we note that one can write

$$\sigma^{*2} = \omega^{*2} - \rho^{*2} \tag{5.4}$$

where

$$\omega^{*2} = \mathcal{E}_{\phi_\Gamma}\left[\left(\frac{\phi_\Gamma}{\phi_\Gamma + \nu}\right)^2 \Big| \underset{\sim}{g}\right] . \tag{5.5}$$

In tabulating constants for the normal approximation, it turns out (Wang, 1973) to be more convenient to tabulate ρ^{*2} and ω^{*2} than ρ^{*2} and σ^{*2} as a function of the prior and sample estimates of ϕ_Γ (λ/ν and $\sum_i (g_i - g.)^2/m$, respectively) given fixed values for m and n . This is so because σ^{*2} is not monotone in the arguments but ω^{*2} is monotone.

6. Numerical Examples

In addition to performing all the computations outlined in sections 2 and 5, the Fortran program MARPRO also provides the joint modal estimates $\tilde{\gamma}_j$ for the gammas. The estimates $\tilde{\pi}_j$ and $\overset{*}{\pi}_j$ of the proportions π_j based on $\tilde{\gamma}_j$ and the marginal mean estimates $\tilde{\tilde{\gamma}}_j$, respectively, are both available from MARPRO for comparisons. Hence, the program MARPRO is recommended for analyzing m-group binomial data with equal n . This program uses the Freeman-Tukey (1950) transformations for binomial data; i.e.,

$$g = \tfrac{1}{2}\left(\sin^{-1}\sqrt{\frac{x}{n+1}} + \sin^{-1}\sqrt{\frac{x+1}{n+1}}\right) , \qquad (6.1)$$

where x is the observed number of successes. In accord with this transformation (6.1), the proportions π_j can be estimated by:

$$\tilde{\pi}_j = \left(1 + \frac{1}{2n}\right)\sin^2\tilde{\gamma}_j - \frac{1}{4n} \quad . \qquad (6.2)$$

(See Novick, Lewis, and Jackson, 1973). Note that $\tilde{\tilde{\pi}}_j$ is also obtained from $\tilde{\tilde{\gamma}}_j$ by equation (6.2).

With the help of this program, we were able to reanalyze the data presented in Table VI of Novick, Lewis, and Jackson (1973). These data were collected for the estimation of item difficulties for six social studies items. For a comparable analysis, we chose to set $\nu = 8$ and $t = 6$ (which is equivalent to let $\lambda = \frac{\nu - 2}{4(t+1)} \simeq .214$ in the current program). In Table 2, estimates of these item difficulties π_j based on $\tilde{\gamma}_j$ and $\overset{*}{\gamma}_j$ were presented. For the joint modal estimates, both the present results (labeled FT) and those of the previous analysis (labeled B, following the cited source) were given. Notice that for some groups slight discrepancies between these two values were found due to different transformations employed [in Novick, Lewis, and Jackson, Anscombe's (1948) transformations were taken]. Both $\rho(= .8856)$ and $\rho^*(= .8906)$ are quite big because of the fairly large sample

sizes (n = 57) for these data. Accordingly, no substantial regressions of individual π_j estimates towards a common value ($\pi.$ corresponding to $\gamma.$) were expected. Similarly, for this large n, the marginal mean estimates do not differ significantly from the joint modal estimates.

In the present analysis, the probability that the item difficulty π_j is greater than some specified value π_o was computed from the marginal distribution of γ_j . The exact (posterior) probabilities and their normal approximations (given in parenthesis), for π_o = .95(-.05).50, are presented in Table 3. The normal approximations were excellent in this case. Thus, having the marginal distributions available, we can now make explicit probability statements about the item difficulties of these six items. For example, one finds the probability that the item difficulty of item 1 is greater than .85 is .9616. These statements should prove to be useful in selecting items for a test. It is interesting to note that the posterior distribution for item one assigns a probability of only .18 to the event $\pi_1 > .95$, even though the observed proportion was .9474. On the other hand, the probability that $\pi_1 < .90$ is .28. Thus, we see that the posterior distribution of π_1 is highly asymmetric, (note that posterior marginal mean estimate of π_1 is .925` in contrast to the posterior distribution of γ_1 which is quite symmetric. For reference, the descriptive statistics (mean, standard deviation, and index of skewness) for the marginal distributions of gamma are also provided in the same table. We noted earlier that a uniform distribution on μ_Γ had been assumed in the derivation when in fact μ_Γ is restricted to the range zero to π/ν . To demonstrate that this does not materially affect the analysis we numerically computed, the aposteriori probability that μ_Γ lie in the range 0 to 2π for each of the data sets presented here. In each instance that probability was unity with an accuracy of 10^{-5}. The point, of course, is that provided m is moderate the prior distribution on μ_Γ will have little effect on the results of the analysis.

To illustrate the differences between the marginal mean and the joint modal estimates when the sample size n is smaller, our second example involves some artificial data sets. Binomial data of m groups were randomly generated from a normal distribution of gamma with mean $\mu_\Gamma = \sin^{-1}\sqrt{\mu_\pi}$ and variance ϕ_Γ. First, m values of g_j were generated by randomly sampling γ_j from the specified normal distribution. These g_j were mapped into p_j by sine-squared transformations ($p_j = \sin^2 g_j$). Then the nearest integers of np_j were taken as the observed number of successes x_j to be analyzed by MARPRO. All the analyses reported hereafter adopt $\nu = 8$ and $\lambda = .25$ (which is equivalent to a value of $t = 5$) for the prior inverse chi-square density of ϕ_Γ.

We have thus generated nine sets (for m = 10(5)50 and n = 8) of data. The values $\mu_\Gamma = 1.1731$, which matches an average of the proportions $\mu_\pi = .85$, and $\phi_\Gamma = .029$ which happens to be $(4n + 2)^{-1}$ for n = 8, were used. Each data set was processed by MARPRO. The results demonstrate consistent patterns for all data sets and with only minor differences for the different values of m . We, therefore, chose to report only the results for m = 10, 20, 30, 40, and 50.

In Table 4, the estimates $\tilde{\tilde{\pi}}_j$, based on the marginal mean estimates $\tilde{\tilde{\gamma}}_j$ and $\tilde{\pi}_j$, (given in parenthesis) based on the joint modal estimates $\tilde{\gamma}_j$ were presented. Since there were many groups having the same observed number of successes x_j and thus, sharing the same estimates of π_j, we present these estimates $\tilde{\pi}_j$ and $\tilde{\tilde{\pi}}_j$ for different values of x instead of for each group. The analyses of these generated data invariably result in significantly bigger values for ρ^* than ρ, so that the general conclusion (3.3) follows. It is also seen that there are substantial differences between $\tilde{\pi}_j$ and $\tilde{\tilde{\pi}}_j$.

As the number of groups (m) increases (for fixed n), both ρ and ρ^* decrease. This means that the estimates of γ_j are more regressed when more groups are used. However, one should bear in mind that sampling fluctuations in these generated data result in small variances in this trend. It was also found that the decreasing rate of ρ^* as m increases is not as high as that of ρ . This confirms the expectation that the joint modal estimates are subject to more influences from other groups. On the other hand, the marginal mean estimates, associated with a squared-error loss for each group separately, place more emphasis on the individual observations. Thus, they are less affected by the inclusion of more groups.

For the marginal probabilities Prob($\pi_j > \pi_o|\underset{\sim}{g}$), we arbitrarily selected those for groups with the number of successes $x_j = 5$ and 7 to be reported in Tables 5 and 6. The values of π_o from .70 to .95 by steps of .05 were included in the tables. The normal approximations are again sufficiently precise. The trend of increases in the probabilities as m increases is consistent with the results in Table 4. Since the marginal distributions are relatively stable w.r.t. the size of m, we suggest that the observed differences are largely due to sampling fluctuations in our generated data. In passing, we note that other data sets generated in the same way described earlier for n = 6 and various sizes of m have also been analyzed. The results reveal the same patterns found in the above example.

Our last example used the result of a 12-item test administered to 35 children. The outcome was that 11 persons scored perfectly, 12 persons missed only one item, and 3, 4, and 5 persons gave correct answers to 8, 9, and 10 items, respectively. The estimates $\tilde{\pi}_j$ and $\tilde{\tilde{\pi}}_j$ and the posterior marginal probabilities prob($\pi_j > \pi_o|\underset{\sim}{g}$) were presented in Table 7. Again, considerable differences between $\tilde{\pi}_j$ and $\tilde{\tilde{\pi}}_j$ were recorded. The posterior

probabilities enable us to reach a more specific judgment on the individual's performance. For example, a person j having 8 correct answers (p_j = .6667) in the test is considered to have an ability greater than .55 with high certainty [$\text{prob}(\pi_j > .55|\underset{\sim}{g}) = .9912$].

The analysis in Table 7 demonstrates the force of the Bayesian m-group method. Consider a class performance as indicated in the data for Table 7 and a situation in which a mastery level of π_o = .85 seemed appropriate. We note that a person answering 10 items correctly has a p_j "score" of .833, and hence, has failed the π_o criterion value of .85. As a result, we would not pass the person. The Bayesian analysis, however, yields a different picture. First, the point estimates of his π are .8829 and .8657 relative to joint zero-one loss and either joint or component squared-error loss, respectively. Thus, on an informal basis, we would probably decide to pass the person. Secondly, the probability that his score is at least .85 is .5082. Therefore, with roughly equal losses associated with false positives and false negatives it would essentially be a toss-up as to whether he was passed or not.

In passing, we also note that for a person j with 11 correct answers, the joint estimate $\tilde{\pi}_j$ is identical to the marginal estimate $\tilde{\tilde{\pi}}_j$ (= .9036). This is so because, for this person, his observed g score (g_j = 1.2288) is equal to the average g score over all persons ($g.$ = 1.2287). It is also clear that, from equation (2.9), the posterior marginal distribution of the corresponding γ_j is symmetric (i.e., the coefficient of skewness δ_j = 0).

Finally, posterior conditional means and standard deviations of γ_j given the marginal mode $\tilde{\tilde{\phi}}_\Gamma$ of $\phi_\Gamma|\underset{\sim}{g}$ were computed for the data of the six social studies items and some of the randomly generated data sets (see Table 8). These conditional mean estimates of γ_j were compared with their marginal

mean estimates to provide some idea about how large the values of m and n would warrant the use of conditional estimates [which are of much simpler form as indicated by equation (2.1)] as approximations to the posterior estimates of γ_j. It appears that for m = 50 and n = 30, the conditional mean estimates $\mathcal{E}(\gamma_j|\tilde{\tilde{\phi}}_\Gamma, g)$ and standard deviations $\sigma(\gamma_j|\tilde{\tilde{\phi}}_\Gamma|g)$ are reasonably close to their marginal statistics $\mathcal{E}(\gamma_j|\underset{\sim}{g})$ and $\sigma(\gamma_j|\underset{\sim}{g})$. Note that referring to equation (2.2), the marginal mode $\tilde{\tilde{\phi}}_\Gamma$ of ϕ_Γ given $\underset{\sim}{g}$ can be obtained by solving the following cubic equation for $\tilde{\tilde{\phi}}_\Gamma$:

$$(m + \nu + 1)\ \tilde{\tilde{\phi}}_\Gamma^{\,3} + [(m + 2\nu + 3)v - \sum_i(g_i - g.)^2 - \lambda]\ \tilde{\tilde{\phi}}_\Gamma^{\,2}$$

$$+ [(\nu + 2)v^2 - 2\lambda v]\ \tilde{\tilde{\phi}}_\Gamma - \lambda v^2 = 0 \quad . \tag{6.3}$$

7. Summary and Conclusions

The knowledge of the posterior marginal distributions of gammas should aid in making our decision when it is concerned about individual persons. There is little doubt that the normal approximations to these marginal distributions are very successful, judging from comparisons with the exact probabilities obtained by integrations. Thus, we recall from section 5 that once $\rho^* = \mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma}{\phi_\Gamma + v}\middle|\underset{\sim}{g}\right)$ and $\sigma^{*2} = \underset{\phi_\Gamma}{\mathrm{Var}}\left(\frac{\phi_\Gamma}{\phi_\Gamma + v}\middle|\underset{\sim}{g}\right)$ are computed by integrations, the interesting descriptive statistics (mean, standard deviation) for γ_j given $\underset{\sim}{g}$ are readily available. Moreover, given ρ^* and σ^{*2}, the relevant probabilities for making individual decisions and the percentage points for γ_j (or π_j) given $\underset{\sim}{g}$ can be satisfactorily approximated using a standard normal table.

Table 1: An Example of the Posterior Densities and Cumulative Probabilities for γ_j given $\underset{\sim}{g}$

c	Posterior Density		Cumulative Probability	
	Exact	Normal Approximation	Exact	Normal Approximation
-2.5	.1834	.1736	.0071	.0062
-2.0	.5346	.5346	.0239	.0228
-1.5	1.2577	1.2825	.0674	.0668
-1.0	2.3558	2.3961	.1575	.1587
- .5	3.4708	3.4863	.3058	.3085
0.0	3.9806	3.9505	.4976	.5000
.5	3.5292	3.4863	.6911	.6915
1.0	2.4118	2.3961	.8426	.8413
1.5	1.2726	1.2825	.9345	.9332
2.0	.5215	.5346	.9779	.9772
2.5	.1676	.1736	.9939	.9938

The prior distribution of ϕ_Γ in this analysis is an inverse chi-square with d.f. $\nu = 8$ and parameter $\lambda = .25$. There are 35 persons, 12 observations in each. The number of persons having 8, 9, 10, 11, and 12 successes are, respectively, 3, 4, 5, 12, and 11. The value g_j corresponding to 10 successes is $g_j = 1.1187$ and the mean of g_j over 35 persons is $g. = 1.2287$. The weight for the modal estimates of $\underset{\sim}{\gamma}$ is $\rho = .2757$ and the weight for the marginal mean estimates is $\rho^* = .4920$. The descriptive statistics of the distribution for γ_j given $\underset{\sim}{g}$ are: $\mu_j = \mathcal{E}(\gamma_j|\underset{\sim}{g}) = 1.1746$, $\sigma_j = [\mathrm{Var}(\gamma_j|\underset{\sim}{g})]^{\frac{1}{2}} = .1010$ and δ_j = coefficient of skewness = -.0035.

Table 2: An Analysis of Item Difficulties for Six Social Studies Items

Item Number	n	p_j	Joint Est. ($\tilde{\pi}_j$) FT	Joint Est. ($\tilde{\pi}_j$) B	Marginal Est. $\tilde{\tilde{\pi}}_j$
1	57	.947	.924	.922	.925
2	57	.386	.423	.423	.421
3	57	.526	.546	.546	.546
4	57	.842	.825	.823	.825
5	57	.772	.762	.761	.762
6	57	.614	.623	.622	.623

Prior distribution of ϕ_T: $\nu = 8$, $t = 6$ (equivalently, $\lambda = .214$); $\rho = .8856$ and $\rho^* = .8906$.

Table 3: Posterior Probabilities $\text{prob}(\pi_j > \pi_o | g)$ for the Six Social Studies Items

Items / π_o	prob($\pi_j > \pi_o \| g$) 1	2	3	4	5	6
.95	.1802 (.1803)	----	----	.0005 (.0005)	0. (0.)	----
.90	.7193 (.7200)	----	----	.0372 (.0369)	.0013 (.0012)	----
.85	.9616 (.9611)	----	----	.2784 (.2789)	.0347 (.0346)	0. (0.)
.80	.9975 (.9974)	----	0. (0.)	.6757 (.6762)	.2205 (.2207)	.0008 (.0008)
.75	.9999 (.9999)	----	.0003 (.0003)	.9203 (.9201)	.5722 (.5725)	.0135 (.0136)
.70	1.0 (1.0)	0. (0.)	.0054 (.0056)	.9893 (.9890)	.8585 (.8584)	.0937 (.0939)
.65	----	.0001 (.0002)	.0448 (.0452)	.9992 (.9991)	.9728 (.9727)	.3198 (.3196)
.60	----	.0024 (.0026)	.1903 (.1902)	1.0 (1.0)	.9970 (.9970)	.6379 (.6375)
.55	----	.0220 (.0225)	.4700 (.4692)	----	.9998 (.9998)	.8760 (.8760)
.50	----	.1101 (.1104)	.7629 (.7626)	----	.9999 (.9999)	.9743 (.9745)
$\tilde{\pi}_j$	.9250	.4211	.5455	.8254	.7622	.6227
p_j	.9474	.3860	.5263	.8421	.7719	.6140
g_j	1.3232	.6724	.8112	1.1543	1.0674	.8984
$\mathcal{E}(\gamma_j \| g)$	1.2865	.7069	.8306	1.1360	1.0587	.9083
$\sigma(\gamma_j \| g)$	.0643	.0641	.0633	.0632	.0629	.0630
δ_j	.0011	-.0010	-.0007	.0007	.0004	-.0004

The exact probabilities were obtained by numerical integrations. Their corresponding normal approximations were given in parentheses. Those probabilities less than .0001 and greater than .9999 were regarded as 0 and 1, respectively, p_j is the observed sampled proportion and δ_j is the index of skewness of the conditional distribution of $\gamma_j | g$. For this set of data, $g. = .9878$ and $\sum_i (g_i - g.)^2 = .2852$.

Table 4: Analyses of Data Sets Randomly Generated

estimates of π_j

x_j \ m	10	20	30	40	50
3	----	.632	----	----	.672
(p_j = .375)	----	(.713)	----	----	(.776)
4	----	.685	----	.718	.715
(p_j = .500)	----	(.743)	----	(.785)	(.791)
5	.720	.735	.735	.759	.756
(p_j = .625)	(.740)	(.772)	(.772)	(.802)	(.805)
6	.772	.786	.781	.800	.798
(p_j = .750)	(.779)	(.802)	(.794)	(.819)	(.820)
7	.828	.841	.829	.844	.843
(p_j = .875)	(.821)	(.835)	(.819)	(.839)	(.837)
8	.913	.923	.905	.913	.913
(p_j = 1.000)	(.889)	(.888)	(.861)	(.871)	(.865)
ρ^*	.4620	.4603	.4068	.3789	.3853
ρ	(.3518)	(.2792)	(.2079)	(.1679)	(.1446)
g.	1.0853	1.1146	1.0950	1.1260	1.1238
$\sum_i (g_i - g.)^2$	.1723	.7388	.8257	1.0257	1.5945

These data were randomly generated from a normal distribution for γ with mean $\mu_\Gamma = 1.1731$ ($\mu_\pi = .85$) and $\phi_\Gamma = .029$. The number of observations in each group is n = 8. The present analyses adopt $\nu = 8$, $t = 5$ (equivalently $\lambda = .25$) for the prior inverse chi-square density of ϕ_Γ . Marginal estimates $\tilde{\tilde{\pi}}_j$ and joint estimates $\tilde{\pi}_j$ (given in parentheses) are presented here. Blank entries indicate there are no values of the corresponding x_j being sampled. $p_j = x_j/n$ is the observed sample proportion of group j . Note that ρ^*_{50} (.3853) for m = 50 is larger than ρ^*_{40} (.3789) for m = 40 due to sampling fluctuations. The generated data for m = 50 has a bigger mean squared deviations of g $[\sum_i (g_i - g.)^2/m]$ than that of the data for m = 40 (.03189 as compared to .02564).

Table 5: Posterior Marginal Distributions of γ_j for Groups with 5 Successes in 8 Trials for the Five Generated Data Sets

$\text{prob}(\pi_j > \pi_o | g)$

π_o \ m	10	20	30	40	50
.95	.0024 (.0027)	.0028 (.0032)	.0016 (.0017)	.0022 (.0025)	.0021 (.0024)
.90	.0207 (.0222)	.0248 (.0268)	.0179 (.0192)	.0254 (.0273)	.0244 (.0261)
.85	.0783 (.0809)	.0933 (.0961)	.0783 (.0809)	.1077 (.1102)	.1037 (.1062)
.80	.1912 (.1923)	.2233 (.2236)	.2063 (.2072)	.2683 (.2677)	.2604 (.2600)
.75	.3514 (.3590)	.3991 (.3952)	.3906 (.3876)	.4773 (.4725)	.4667 (.4623)
.70	.5299 (.5248)	.5839 (.5778)	.5886 (.5832)	.6773 (.6720)	.6672 (.6621)
$\tilde{\pi}_j$	.720	.735	.735	.759	.756
$\mathcal{E}(\gamma_j \| g)$	.9989	1.0149	1.0149	1.0397	1.0369
$\sigma(\gamma_j \| g)$	.1245	.1213	.1130	.1089	.1093
δ_j	-.0091	-.0099	-.0080	-.0083	-.0079

For these groups, $p_j = .625$ and $g_j = .8982$. The exact probabilities $\text{prob}(\pi_j > \pi_o | g)$ and the corresponding normal approximations (in parentheses) are presented.

Table 6: Posterior Marginal Distributions of γ_j for Groups with 7 Successes in the Five Generated Data Sets

prob($\pi_j > \pi_o | g$)

π_o \ m	10	20	30	40	50
.95	.0335 (.0325)	.0390 (.0384)	.0230 (.0220)	.0273 (.0267)	.0271 (.0265)
.90	.1421 (.1435)	.1652 (.1671)	.1228 (.1241)	.1477 (.1494)	.1463 (.1479)
.85	.3238 (.3268)	.3678 (.3701)	.3137 (.3165)	.3673 (.3693)	.3638 (.3658)
.80	.5327 (.5342)	.5866 (.5867)	.5432 (.5441)	.6102 (.6096)	.6056 (.6051)
.75	.7171 (.7163)	.7658 (.7641)	.7421 (.7407)	.8000 (.7981)	.7959 (.7941)
.70	.8493 (.8476)	.8841 (.8825)	.8758 (.8740)	.9135 (.9123)	.9109 (.9097)
$\tilde{\pi}_j$	.828	.841	.829	.844	.843
$\mathcal{E}(\gamma_j \| g)$	1.1177	1.1334	1.1196	1.1371	1.1360
$\sigma(\gamma_j \| g)$	.1234	.1198	.1121	.1077	.1082
δ_j	.0034	.0019	.0025	.0011	.0011

For these groups, $p_j = .875$ and $g_j = 1.1554$.

Table 7: Analyses of a Data Set Obtained from a 12-item Test Given to 35 Children

π_o	$x_j = 8$ $p_j = .6667$ $g_j = .9423$	$x_j = 9$ $p_j = .7500$ $g_j = 1.0262$	$x_j = 10$ $p_j = .8333$ $g_j = 1.1187$	$x_j = 11$ $p_j = .9167$ $g_j = 1.2288$	$x_j = 12$ $p_j = 1.0000$ $g_j = 1.4303$
	$\text{prob}(\pi_j \geq \pi_o \mid g)$				
.95	.0052(.0062)	.0156(.0168)	.0444(.0455)	.1224(.1234)	.4275(.4321)
.90	.0557(.0585)	.1174(.1192)	.2298(.2304)	.4193(.4199)	.7808(.7807)
.85	.2035(.2035)	.3348(.3327)	.5082(.5058)	.7110(.7097)	.9379(.9359)
.80	.4304(.4253)	.5898(.5853)	.7502(.7478)	.8873(.8864)	.9861(.9850)
.75	.6588(.6534)	.7917(.7894)	.8966(.8964)	.9642(.9643)	.9974(.9971)
.70	.8276(.8262)	.9115(.9123)	.9642(.9653)	.9905(.9909)	-- (--)
.65	.9258(.9276)	.9680(.9699)	.9895(.9905)	-- (--)	-- (--)
.60	.9725(.9750)	.9901(.9915)	.9973(.9979)	-- (--)	-- (--)
.55	.9912(.9929)	-- (--)	-- (--)	-- (--)	-- (--)
	Estimates of Proportions and Descriptive Statistics				
$\tilde{\pi}_j$	(.8468)	(.8644)	(.8829)	(.9036)	(.9376)
$\tilde{\tilde{\pi}}_j$	.7961	.8305	.8657	.9036	.9606
$\mathcal{E}(\gamma_j \mid g)$	1.0878	1.1291	1.1746	1.2287	1.3279
$\sigma(\gamma_j \mid g)$	.1029	.1018	.1010	.1007	.1018
δ_j	-.0089	-.0064	-.0035	.0000	.0063

For these data, $g. = 1.2287$, $\sum_i (g_i - g.)^2 = .9175$. Prior distribution of ϕ_T: $\nu = 8$, $t = 5$ (equivalently, $\lambda = .25$). $\rho = .2757$ and $\rho^* = .4920$. The figures in parentheses are normal approximations to $\text{prob}(\pi_j \geq \pi_o \mid g)$. Those cumulative probabilities greater than .999 were omitted in the table.

$x_j = 8$ prob$(.8861 < \gamma_j < 1.2895) = .95$ or prob$(.6000 < \pi_j < .9229) = .95$

$x_j = 9$ prob$(.9296 < \gamma_j < 1.3286) = .95$ or prob$(.6422 < \pi_j < .9425) = .95$

$x_j = 10$ prob$(.9766 < \gamma_j < 1.3726) = .95$ or prob$(.6866 < \pi_j < .9612) = .95$

$x_j = 11$ prob$(1.0313 < \gamma_j < 1.4261) = .95$ or prob$(.7361 < \pi_j < .9792) = .95$

$x_j = 12$ prob$(1.1284 < \gamma_j < 1.5274) = .95$ or prob$(.8167 < \pi_j < .9981) = .95$

Table 8: Conditional Distributions of γ_j Given the Marginal Mode $\tilde{\phi}_\Gamma$ of $\phi_\Gamma|g$

I. The Six Social Studies Items (m = 6, n = 57)

$\tilde{\phi}_\Gamma$ = .03004

p_j	.9474	.3860	.5263	.8421	.7719	.6140
$\tilde{\pi}_j$	.9250	.4211	.5455	.9254	.7622	.6227
$\tilde{\pi}_j\|\tilde{\phi}_\Gamma$	(.9219)	(.4265)	(.5486)	(.8233)	(.7610)	(.6242)
$\mathcal{E}(\gamma_j\|g)$	1.2865	.7069	.8306	1.1360	1.0857	.9083
$\mathcal{E}(\gamma_j\|\tilde{\phi}_\Gamma, g)$	(1.2808)	(.7122)	(.8336)	(1.1332)	(1.0574)	(.9097)
$\sigma(\gamma_j\|g)$	.0643	.0641	.0633	.0632	.0629	.0630
$\sigma(\gamma_j\|\tilde{\phi}_\Gamma, g)$*	(.0624)	(.0624)	(.0624)	(.0624)	(.0624)	(.0624)

II. Randomly Generated Data (m = 50, n = 8)

$\tilde{\phi}_\Gamma$ = .01652

p_j	.3750	.5000	.6250	.7500	.8750	1.000
$\tilde{\pi}_j$	.6717	.7147	.7561	.7977	.8427	.9129
$\tilde{\pi}_j\|\tilde{\phi}_\Gamma$	(.6833)	(.7231)	(.7614)	(.8002)	(.8421)	(.9081)
$\mathcal{E}(\gamma_j\|g)$	.9499	.9934	1.0369	1.0829	1.1360	1.2306
$\mathcal{E}(\gamma_j\|\tilde{\phi}_\Gamma, g)$	(.9615)	(1.0021)	(1.0427)	(1.0856)	(1.1352)	(1.2234)
$\sigma(\gamma_j\|g)$	.1125	.1106	.1093	.1084	.1082	.1098
$\sigma(\gamma_j\|\tilde{\phi}_\Gamma, g)$*	(.1047)	(.1047)	(.1047)	(.1047)	(.1047)	(.1047)

III. Randomly Generated Data ($m = 50$, $n = 30$)

$\tilde{\phi}_\Gamma = .00924$

p_j	.6667	.7667	.8000	.8333	.8667	.9000	.9333	.9667
$\tilde{\pi}_j$	.7535	.8031	.8200	.8374	.8552	.8740	.8941	.9167
$\tilde{\tilde{\pi}}_j \mid \tilde{\phi}_\Gamma$	(.7564)	(.8045)	(.8209)	(.8377)	(.8550)	(.8733)	(.8928)	(.9150)
$\mathcal{E}(\gamma_j \mid g)$	1.0464	1.1048	1.1259	1.1483	1.1723	1.1987	1.2289	1.2660
$\mathcal{E}(\gamma_j \mid \tilde{\phi}_\Gamma, g)$	(1.0498)	(1.1065)	(1.1270)	(1.1487)	(1.1720)	(1.1977)	(1.2270)	(1.2630)
$\sigma(\gamma_j \mid g)$	.0685	.0677	.0675	.0675	.0674	.0675	.0678	.0683
$\sigma(\gamma_j \mid \tilde{\phi}_\Gamma, g)$*	(.0665)	(.0665)	(.0665)	(.0665)	(.0665)	(.0665)	(.0665)	(.0665)

Data sets I and II were used in Tables 2 and 4, respectively. Data set III was generated specifically for this table. The sample statistics for Data set III are $g. = 1.1625$ and $\Sigma(g_i - g.)^2 = .2927$.

*The conditional standard deviations of γ_j given ϕ_Γ and g are same for all groups.

References

Anscombe, F. T. The transformation of Poisson, binomial, and negative binomial data. Biometrika, 1948, 35, 246-254.

Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: McGraw-Hill, 1972.

De Finetti, B. Foresight: its logical laws, its subjective sources. In H. E. Kyburg, Jr., and G. E. Smokler (Eds.), Studies in Subjective Probability. New York: Wiley, 1964, 93-105. (The original French version appeared in Ann. de L'Inst. Henri Poincaré, 1937. 7.)

Freeman, M. F., and Tukey, J. W. Transformations related to the angular and the square root. The Annals of Mathematical Statistics, 1950, 21, 607-611.

Hambleton, R. K., and Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, in press.

Hill, B. M. Inference about variance components in the one-way model. Journal of the American Statistical Association, 1965, 60, 806-825.

Jackson, P. H. Some simple approximations in the estimation of many parameters. British Journal of Mathematical and Statistical Psychology, 1972, 25, 213-228.

Kelley, T. L. Statistical Method. New York: MacMillan, 1923.

Kelley, T. L. Interpretation of Educational Measurements. Yonkers on Hudson, New York: World Book, 1927.

Leonard, T. Bayesian methods for binomial data: Some computational methods and further ideas. ACT Technical Bulletin No. 7. Iowa City, Iowa: The American College Testing Program, 1972.

Lindley, D. V. and Smith, A. F. M. Bayes estimates for the linear model. Journal of the Royal Statistical Society, Series B, 1972, 34, 1-41.

Novick, M. R. Bayesian considerations in educational information systems. In G. V, Glass (Ed.), Proceedings of the 1970 Invitational Conference on Testing Problems. Princeton, New Jersey: Educational Testing Service, 1970.

Novick, M. R., & Jackson, P. H. Bayesian guidance technology. The Review of Educational Research, 1970, 40(4), 459-494.

Novick, M. R. and Jackson, P. H. Statistical Methods for Educational and Psychological Research, New York: McGraw-Hill, 1974.

Novick, M. R. Lewis, C., and Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-45.

Ralston, A. A First Course in Numerical Analysis. New York: McGraw-Hill, 1965.

Wang, M. M. Tables of constants for the posterior marginal estimates of proportions in m groups. ACT Technical Bulletin No. 14. Iowa City, Iowa: The American College Testing Program, 1973.

ACT TECHNICAL BULLETIN

SUPPLEMENT NO. 13-1

A PROPER PRIOR FOR μ_r IN ESTIMATING PROPORTIONS OF m GROUPS*

by

Charles Lewis
The University of Illinois

Ming-mei Wang
The American College Testing Program

and

Melvin R. Novick

The American College Testing Program
and
The University of Iowa

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

October, 1973

*The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with The Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

A Proper Prior for μ_Γ in Estimating Proportions of m Groups*

by

Charles Lewis
The University of Illinois

Ming-mei Wang
The American College Testing Program

and

Melvin R. Novick
The American College Testing Program
and
The University of Iowa

Before concluding our discussion of the estimation of proportions in m groups, we shall briefly consider the effect of assuming a proper prior distribution for μ_Γ, in place of the uniform distribution used up until now. The form of the prior is specified, if we restrict ourselves to natural conjugate densities, by

$$b(\mu_\Gamma|\phi_\Gamma) \propto \phi_\Gamma^{-\frac{1}{2}} \exp[-\tfrac{1}{2} k \phi_\Gamma^{-1}(\mu_\Gamma - h)^2] , \tag{1}$$

where h is the prior mean for μ_Γ and k the "prior sample size" associated with our knowledge of μ_Γ . Combining Equation (1) with $b(\underset{\sim}{\gamma}|\mu_\Gamma, \phi_\Gamma)$ and $b(\phi_\Gamma)$, we obtain

$$b(\underset{\sim}{\gamma}, \mu_\Gamma, \phi_\Gamma) \propto \phi_\Gamma^{-(\nu + m + 3)} \exp\{-\tfrac{1}{2}\phi_\Gamma^{-1}[\Sigma(\gamma_i - \mu_\Gamma)^2 + k(\mu_\Gamma - h)^2 + \lambda]\} . \tag{2}$$

*This note is a Technical Supplement to ACT Technical Bulletin No. 13. The material contained here should be considered as inserted prior to the concluding section of that Bulletin.

Now, with some rearranging of terms, we may write

$$\phi_\Gamma^{-1}[\Sigma(\gamma_i - \mu_\Gamma)^2 + k(\mu_\Gamma - h)^2 + \lambda]$$

$$= (k + m)\phi_\Gamma^{-1}(\mu_\Gamma - \frac{kh + m\gamma.}{k + m})^2$$

$$+ \phi_\Gamma^{-1}[\lambda + \Sigma(\gamma_i - \gamma.)^2 + \frac{km}{k + m}(h - \gamma.)^2] .$$

Thus, if we integrate Equation (2) w.r.t. μ_Γ, we obtain

$$b(\underset{\sim}{\gamma}, \phi_\Gamma) \propto \phi_\Gamma^{-\frac{1}{2}(\nu + m + 2)} \exp\{-\tfrac{1}{2}\phi_\Gamma^{-1}[\lambda + \Sigma(\gamma_i - \gamma.)^2 + \frac{km}{k + m}(h - \gamma.)^2]\} .$$

Further integration, this time w.r.t. ϕ_Γ, yields

$$b(\underset{\sim}{\gamma}) \propto [\lambda + \Sigma(\gamma_i - \gamma.)^2 + \frac{km}{k + m}(h - \gamma.)^2]^{-\frac{1}{2}(\nu + m)}$$

$$\propto [\nu + (\underset{\sim}{\gamma} - h\underset{\sim}{1})'\underset{\sim}{A}(\underset{\sim}{\gamma} - h\underset{\sim}{1})]^{-\frac{1}{2}(\nu + m)} , \qquad (3)$$

where $\underset{\sim}{1}$ is the vector of order m all of whose elements are unity and $\underset{\sim}{A}$ is the m x m matrix with diagonal elements $(\frac{k + m - 1}{k + m})\frac{\nu}{\lambda}$ and off-diagonal elements $(\frac{-1}{k + m})\frac{\nu}{\lambda}$. In other words, we have shown that the unconditional prior distribution for γ is multivariate t, with ν degrees of freedom, mean $h\underset{\sim}{1}$, and covariance matrix $(\frac{\nu}{\nu - 2})\underset{\sim}{A}^{-1}$. In particular, this implies that the marginal prior density of any γ_i is univariate t, with ν degrees of freedom, mean h, and variance $(\frac{k + 1}{k})\frac{\lambda}{\nu - 2}$, provided k is greater than zero. We note that if k = 0, the joint density $b(\underset{\sim}{\gamma})$ in Equation (3) becomes improper because the inverse of $A = I_m - \frac{\underset{\sim}{1}\underset{\sim}{1}'}{m}$ does not exist.

Novick, Lewis, and Jackson (1973) discuss the possibility of interrogating an investigator about his prior beliefs concerning π_i, where i has been arbitrarily selected. One of their suggestions is to approximate these

beliefs with a beta density. If we interpret the parameters of the density so obtained as the numbers of "prior successes" and "prior failures", respectively, then the sum of the parameters gives the "prior sample size", t, and the mean of the distribution is the "prior proportion of successes", $M_.$. From these two values, Novick, Lewis, and Jackson (1973) obtain approximate expressions for the mean and variance of $\gamma_i = \sin^{-1}\sqrt{\pi_i}$, namely

$$\mathcal{E}(\gamma_i) \doteq \sin^{-1}\sqrt{M}$$

and

$$\mathrm{Var}(\gamma_i) \doteq \frac{1}{4(t+1)} \quad .$$

If we now equate these values to the mean and variance for γ_i found above, we have expressions for h and for λ

$$h = \sin^{-1}\sqrt{M} \tag{4}$$

and

$$\lambda = \frac{k(\nu-2)}{4(k+1)(t+1)} \, . \tag{5}$$

Novick, Lewis, and Jackson (1973) have argued that $\nu = 8$ will, in many cases, be a reasonable specification of the prior degrees of freedom for ϕ_Γ . If we accept this value, then our only remaining task is to specify k, the "prior sample size" for μ_Γ . It is tempting, and may in some cases be reasonable, to assume that our prior knowledge of μ_Γ and of ϕ_Γ come from essentially the same sources and so could be associated with a single hypothetical prior sample. This would allow us to equate k - 1 and ν, giving a value of k = 9 in the present circumstances. In many cases, however, when we have selected our groups (or individuals) to be quite similar, our knowledge concerning ϕ_Γ may be greater than our knowledge of

μ_Γ . This would suggest taking $k < \nu + 1$. Working with an improper prior for μ_Γ represents, in effect, the extreme situation where $k = 0$. If we were to work with $k = 5$, for instance, Equation (5) would reduce to

$$\lambda = \frac{5}{4(t+1)} ,$$

which may often be a reasonable assignment.

Once values have been supplied for h, k, λ, and ν, we can work directly with the posterior distribution for $\underset{\sim}{\gamma}$, which is proportional to the product of the likelihood $\ell(\underset{\sim}{\gamma}|\underset{\sim}{g})$ and the prior density of $\underset{\sim}{\gamma}$, given in Equation (3):

$$b(\underset{\sim}{\gamma}|\underset{\sim}{g}) \propto \exp[-\tfrac{1}{2}\Sigma v_i^{-1}(g_i - \gamma_i)^2] \cdot [\lambda + \Sigma(\gamma_i - \gamma.)^2 + \frac{km}{k+m}(h - \gamma.)^2]^{-\frac{1}{2}(\nu+m)} . \tag{6}$$

If we take derivatives with respect to each γ_i and set the results equal to zero, we obtain the following equations for the joint posterior mode of $\underset{\sim}{\gamma}$:

$$\tilde{\gamma}_j = \frac{\tilde{\phi}_\Gamma g_j + v_j \tilde{\mu}_\Gamma}{\tilde{\phi}_\Gamma + v_j} , \tag{7}$$

where

$$\tilde{\mu}_\Gamma = \frac{kh + m\tilde{\gamma}.}{k + m} ,$$

and

$$\tilde{\phi}_\Gamma = [\lambda + \Sigma(\tilde{\gamma}_i - \tilde{\gamma}.)^2 + \frac{km}{k+m}(\tilde{\gamma}. - h)^2]/(\nu + m) .$$

These equations are closely related to Equation (3.1), Section 3, of the main text; the solution obtained with an improper prior for μ_Γ . Indeed, if $h = \tilde{\gamma}.$ or if $k = 0$, the two results are identical except for a difference of unity in the denominator of $\tilde{\phi}_\Gamma$. At a practical level, making use of our

prior knowledge about μ_Γ increases the effective number of groups in the study; this will be particularly important in cases where m is relatively small (say, between 5 and 15). On the other hand, for larger m or in cases where the prior specification closely agrees wit' the sample results, there will be little to choose between proper and improper priors for μ_Γ .

References

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-45.

ACT TECHNICAL BULLETIN NO. 14

TABLES OF CONSTANTS FOR THE POSTERIOR MARGINAL ESTIMATES OF PROPORTIONS IN m GROUPS

by

Ming-mei Wang

The American College Testing Program

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

June, 1973

TABLES OF CONSTANTS FOR THE POSTERIOR MARGINAL ESTIMATES OF PROPORTIONS IN m GROUPS

by

Ming-mei Wang

The American College Testing Program

1. General Descriptions

For estimation of proportions in m groups, Novick, Lewis, and Jackson (1973) have developed a Bayesian Model II solution which provides posterior joint modal estimates $\tilde{\gamma}_j$ of the transformed proportions $\gamma_j = \sin^{-1}\sqrt{\pi_j}$. The values $\tilde{\pi}_j$ $(j = 1, \ldots, m)$, the sine-squared transformations of $\tilde{\gamma}_j$ suitably corrected by a factor depending on sample sizes n_j, were then taken as the Bayesian joint estimates of the group proportions π_j . These joint estimates are useful in making joint decisions for m groups.

To aid in making separate decisions on individual groups, the posterior marginal distributions of γ_j, for the case of equal sample sizes n, have been studied recently by Lewis, Wang, and Novick (1973). They worked out the posterior marginal mean estimates $\mu_j = \mathcal{E}(\gamma_j|g)$ of γ_j to be

$$\mu_j = \rho^* g_j + (1 - \rho^*) g. \quad , \tag{1}$$

where $\rho^* = \mathcal{E}_{\phi_r}\left(\frac{\phi_r}{\phi_r + v} \middle| \underset{\sim}{g}\right)$ and $v = (4n + 2)^{-1}$. The posterior variances σ_j^2 of γ_j were expressed as:

$$\sigma_j^2 = \mathrm{Var}(\gamma_j|\underset{\sim}{g}) = v\left[\frac{(m - 1)\ \rho^* + 1}{m}\right] + (g_j - g.)^2\ \sigma^{*2} \ , \tag{2}$$

where $\sigma^{*2} = \mathrm{Var}_{\phi_r}\left(\frac{\phi_r}{\phi_r + v} \middle| \underset{\sim}{g}\right)$.

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U.S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Thus, if ρ^* and σ^{*2} have been computed, one could easily obtain the posterior means and variances of γ_j from formulas (1) and (2).

To compute ρ^* and σ^{*2}, one needs to know the posterior distribution of ϕ_Γ. This was given by Lewis, Wang, and Novick (1973) as

$$b(\phi_\Gamma|\underset{\sim}{g}) \propto (\phi_\Gamma + v)^{-\frac{m-1}{2}} \exp[-\tfrac{1}{2}(\phi_\Gamma + v)^{-1} \Sigma(g_i - g.)^2]$$

$$\cdot\ \phi_\Gamma^{-(\frac{\nu}{2}+1)} \exp(-\tfrac{1}{2}\phi_\Gamma^{-1}\lambda) \ . \qquad (3)$$

The values of ρ^* and σ^{*2} can then be obtained by numerical integrations:

$$\rho^* = \mathcal{E}_{\phi_\Gamma}\left(\frac{\phi_\Gamma}{\phi_\Gamma + v} \,|\underset{\sim}{g}\right)$$

$$= \int_0^\infty \phi_\Gamma(\phi_\Gamma + v)^{-1}\, b(\phi_\Gamma|\underset{\sim}{g})d\phi_\Gamma \ , \qquad (4)$$

and

$$\sigma^{*2} = \omega^{*2} - \rho^{*2} \ , \qquad (5)$$

where

$$\omega^{*2} = \int_0^\infty \phi_\Gamma^{\,2}(\phi_\Gamma + v)^{-2}\, b(\phi_\Gamma|\underset{\sim}{g})d\phi_\Gamma \ . \qquad (6)$$

Equations (3) - (6) indicate that ρ^* and σ^{*2} would vary for different m, n, ν, λ and $S_g^2 = \Sigma(g_i - g.)^2$. Consequently, a complete set of tables of ρ^*, σ^{*2} for all practical values of these five parameters would require a formidable volume. Since previous experiences and theoretical findings have suggested that $\nu = 8$ was a satisfactory choice in most applications for the prior distribution of ϕ_Γ (see Novick, Lewis, and Jackson, 1973), we are, therefore, content with providing a subset of the tables which set $\nu = 8$. The values $m = 10(5)30(10)80$ and $n = 8(2)30$ are included in the tables presented here. For each pair of (m, n), ρ^* and σ^{*2} were computed for different values of λ/ν and $s_g^2 = S_g^2/m$ (prior and sample estimates of ϕ_Γ). Values of λ/ν, $s_g^2 = .01(.01).05$ are included in the tables. For other

values of m, n, λ/ν and s_g^2 within their ranges selected in these tables, corresponding ρ^* and σ^{*2} can be approximated by interpolation.

While σ^{*2} is not monotone in λ/ν and s_g^2, the posterior expectation ω^{*2} of $\phi_\Gamma^2(\phi_\Gamma + v)^{-2}$ with respect to ϕ_Γ is monotone. We, therefore, tabulate ρ^* and ω^{*2} instead of ρ^* and σ^{*2}. Given a prior estimate λ/ν and a sample estimate s_g^2 of ϕ_Γ, for which ρ^* and ω^{*2} are not explicitly tabulated, one can obtain ρ^* and ω^{*2} by interpolation using the .iven tables. The value of σ^{*2} then can be found by subtracting ρ^{*2} from ω^{*2} $(\sigma^{*2} = \omega^{*2} - \rho^{*2})$.

It may be noted that these tables include the size of m and n only up to 80 and 30, respectively. For values of $m > 80$ and $n > 30$, the posterior conditional distribution of γ_j given $\tilde{\tilde{\phi}}_\Gamma$ [the posterior modal estimates of ϕ_Γ obtained from its posterior density expressed by (3)] and $\underset{\sim}{g}$ was found to satisfactorily approximate the posterior marginal distribution of γ_j given $\underset{\sim}{g}$. This posterior conditional distribution of γ_j given $\tilde{\tilde{\phi}}_\Gamma$, $\underset{\sim}{g}$ was shown to be normal (see section 2, Lewis, Wang, and Novick, 1973) with mean

$$\mathcal{E}(\gamma_j|\tilde{\tilde{\phi}}_\Gamma, \underset{\sim}{g}) = \frac{\tilde{\tilde{\phi}}_\Gamma g_j + vg.}{\tilde{\tilde{\phi}}_\Gamma + v} , \tag{7}$$

and variance

$$\mathrm{Var}(\gamma_j|\tilde{\tilde{\phi}}_\Gamma, \underset{\sim}{g}) = \frac{v(\tilde{\tilde{\phi}}_\Gamma + m^{-1}v)}{\tilde{\tilde{\phi}}_\Gamma + v} . \tag{8}$$

Thus, for large m and n, this conditional distribution provides an approximate basis for making decisions on individual groups. Having made ρ^* and σ^{*2} available, the probabilities that a group proportion π_j is greater than some criterion π_o given observed $\underset{\sim}{g}$ [$\mathrm{prob}(\pi_j \geq \pi_o|\underset{\sim}{g})$] can be obtained applying the normal approximation to the posterior distribution of γ_j given $\underset{\sim}{g}$ discussed in (Lewis, Wang, and Novick, 1973). That is,

$$\text{prob}(\pi_j \geq \pi_o | g) = \text{prob}(\gamma_j \geq \gamma_o | \underset{\sim}{g}) \approx \text{prob}(z \geq z_o) \tag{9}$$

where $\gamma_o = \sin^{-1}\sqrt{\pi_o}$,

$$z_o = [\gamma_o - \rho^*(g_j - g.) - g.]\left[\frac{1 + (m - 1)\rho^*}{(4n + 2)m} + (g_j - g.)^2 \sigma^{*2}\right]^{-\frac{1}{2}} . \tag{10}$$

Similarly, approximate 100α percentage points $\pi_{\alpha j}$ of π_j can be computed with the help of a standard normal table. For example,

$$\pi_{\alpha j} = \sin^2 \gamma_{\alpha j} \tag{11}$$

where

$$\gamma_{\alpha j} = \mu_j + z_\alpha \left[\frac{1 + (m - 1)\rho^*}{(4n + 2)m} + (g_j - g.)^2 \sigma^{*2}\right]^{\frac{1}{2}} , \tag{12}$$

and z_α is the 100α percentage point of a standard normal variate.

2. Interpolations

In practical applications, one would not expect s_g^2 of his data to be exactly equal to the tabulated values (s_g^2 = .01(.01).05). Likewise, an investigator may have reason to choose his prior λ/ν other than values included in these tables. In these cases, approximations of ρ^* and ω^{*2} can be obtained by interpolation using available tabular points. For illustrative purposes, we have computed ρ^* and ω^{*2} for s_g^2 = .0169, .0256, .0361, and .0484 with m = 10, ν = 8, λ/ν = .01 and n = 8, 16 by numerical integrations. These exact values of ρ^* and ω^{*2} are then compared with those (ρ_I^*, ω_I^{*2}) obtained by simple linear interpolation. The table presented below shows that the discrepancies between interpolated and exact values are negligible.

Comparison Between Exact and Interpolated Values of ρ^* and ω^{*2}

(m = 10, ν = 8, λ/ν = .01)

	n = 8				n = 16			
s_g^2	.0169	.0256	.0361	.0484	.0169	.0256	.0361	.0484
ρ^* (exact)	.2538	.2654	.2812	.3029	.3960	.4285	.4730	.5300
ρ_I^* (interpolated)	(.2540)	(.2656)	(.2815)	(.3031)	(.3966)	(.4292)	(.4735)	(.5301)
ω^{*2} (exact)	.0718	.0787	.0886	.1030	.1674	.1957	.2373	.2953
ω_I^{*2} (interpolated)	(.0720)	(.0789)	(.0888)	(.1031)	(.1681)	(.1966)	(.2381)	(.2955)
$\sigma^{*2} = \omega^{*2} - \rho^{*2}$	.0072	.0083	.0095	.0113	.0106	.0121	.0136	.0144
$\sigma_I^{*2} = \omega_I^{*2} - \rho_I^{*2}$	(.0075)	(.0084)	(.0096)	(.0112)	(.0108)	(.0124)	(.0139)	(.0145)

It may be noted that in this example, the monotone functions of both ρ^* and ω^{*2} on s_g^2 are slightly postively accelerated. Consequently, the values obtained from linear interpolations consistently overestimate, though negligibly, the exact values as demonstrated in the above table. However,

the characteristic of the monotone functions of ρ^* and ω^{*2} on s_g^2 varies with the values of m, n, ν and λ/ν. For example, given m = 10, ν = 8, λ/ν = .05, and n = 16, the functions of ρ^* and ω^{*2} on s_g^2 become negatively accelerated. Therefore, whether the interpolated value underestimates or overestimates the exact one depends on other parametric values (e.g., m, n, ν and λ/ν) being considered. In general, the discrepancies are very small when linear interpolation over an interval length of .01 is applied in our present problem. Approximations of ρ^* and ω^{*2} for nontabulated values of m, n, and λ/ν can also be obtained satisfactorily by linear interpolations.

3. A Numerical Example

To illustrate the use of these tables in actual data, the example presented in Table 7 of Lewis, Wang, and Novick (1973) was reanalyzed employing these tables. There were 35 children taking a 12-item test. The sufficient sample statistics for our analysis were $m = 35$, $n = 12$, $g. = 1.2287$, and $s_g^2 = \sum_i (g_i - g.)^2/m = .9175/35 = .02621$. The same prior distribution for ϕ_r (namely, $\nu = 8$, $\lambda/\nu = .25/8 = .03125$) was adopted. For notational convenience, we shall denote $\rho^*(m, n, \lambda/\nu, s_g^2)$ and $\omega^{*2}(m, n, \lambda/\nu, s_g^2)$ as the values of ρ^* and ω^{*2} for given m, n, λ/ν and s_g^2.

Using the table for $m = 30$, $n = 12$, we first find $\rho^*(30, 12, .03, .02) = .4688$ and $\rho^*(30, 12, .03, .03) = .5138$, so that $\rho^*(30, 12, .03, .02621)$ can be approximated by interpolating between these two values:

$$\rho^*(30, 12, .03, .02621) \approx .4688 + \frac{(.02621 - .02)}{(.03 - .02)} \cdot (.5138 - .4688) = .4967 .$$

Similarly, interpolate between $\rho^*(30, 12, .04, .02) = .5162$ and $\rho^*(30, 12, .04, .03) = .5566$, we have

$$\rho^*(30, 12, .04, .02621) \approx .5413 .$$

The next step is to interpolate between $\rho^*(30, 12, .03, .02621)$ and $\rho^*(30, 12, .04, .02621)$ to obtain:

$$\rho^*(30, 12, .03125, .02621) \approx \underline{.5023} .$$

Following the same procedure, $\rho^*(40, 12, .03125, .02621)$ was approximated using the table for $m = 40$, $n = 12$:

$$\left.\begin{array}{l}\rho^*(40, 12, .03, .02) = .4449 \\ \rho^*(40, 12, .03, .03) = .4970\end{array}\right\} \text{gives } \rho^*(40, 12, .03, .02621) \approx .4773;$$

$$\left.\begin{array}{l}\rho^*(40, 12, .04, .02) = .4900 \\ \rho^*(40, 12, .04, .03) = .5370\end{array}\right\} \text{gives } \rho^*(40, 12, .04, .02621) \approx .5192;$$

thus, $\rho^*(40, 12, .03125, .02621) \approx \underline{.4825}$.

Finally, we interpolate between m = 30 and m = 40 to approximate the value for m = 35:

$$\rho^*(35,\ 12,\ .03125,\ .02621) \approx \tfrac{1}{2}(.5023 + .4825) = \boxed{.4924}\ .$$

This value is very close to the exact value (.4920) obtained from our previous analysis. In the same way, $\omega^{*2}(35,\ 12,\ .03125,\ .02621)$ can be approximated from available tables. First,

$$\left.\begin{aligned}\omega^{*2}(30,\ 12,\ .03,\ .02) &= .2255\\ \omega^{*2}(30,\ 12,\ .03,\ .03) &= .2702\end{aligned}\right\}\ \text{gives}\ \omega^{*2}(30,\ 12,\ .03,\ .02621) \approx .2533;$$

and

$$\left.\begin{aligned}\omega^{*2}(30,\ 12,\ .04,\ .02) &= .2718\\ \omega^{*2}(30,\ 12,\ .04,\ .03) &= .3153\end{aligned}\right\}\ \text{gives}\ \omega^{*2}(30,\ 12,\ .04,\ .02621) \approx .2988;$$

so that $\omega^{*2}(30,\ 12,\ .03125,\ .02621) \approx \underline{.2590}$.

Secondly,

$$\left.\begin{aligned}\omega^{*2}(40,\ 12,\ .03,\ .02) &= .2028\\ \omega^{*2}(40,\ 12,\ .03,\ .03) &= .2524\end{aligned}\right\}\ \text{gives}\ \omega^{*2}(40,\ 12,\ .03,\ .02621) \approx .2336;$$

and

$$\left.\begin{aligned}\omega^{*2}(40,\ 12,\ .04,\ .02) &= .2447\\ \omega^{*2}(40,\ 12,\ .04,\ .03) &= .2932\end{aligned}\right\}\ \text{gives}\ \omega^{*2}(40,\ 12,\ .04,\ .02621) \approx .2932;$$

so that $\omega^{*2}(40,\ 12,\ .03125,\ .02621) \approx \underline{.2388}$.

Finally, we arrive at

$$\omega^{*2}(35,\ 12,\ .03125,\ .02621) \approx \tfrac{1}{2}(.2590 + .2388) = \boxed{.2489}.$$

Thus, the approximate values of ρ^* and σ^{*2} for the present data have been obtained:

$$\rho^* = .4924$$

and

$$\sigma^{*2} = \omega^{*2} - \rho^{*2} = .0064\ .$$

Now applying formulas (1) and (2), the posterior marginal mean estimates (μ_j) of γ_j (thus $\tilde{\tilde{\pi}}_j$ of π_j) given $\underset{\sim}{g}$ and the corresponding posterior variances σ_j^2 (or standard deviations σ_j) can easily be computed. The results obtained from the present analysis are compared with the previous results produced by the program MARPRO described in Lewis, Wang, Novick, 1973. In the table presented below, estimates from the present approximate method are given together with those exact estimates (enclosed in parentheses) obtained from MARPRO output.

Posterior Marginal Estimates of γ_j, π_j, and σ_j

	$x_j = 8$	$x_j = 9$	$x_j = 10$	$x_j = 11$	$x_j = 12$
$\tilde{\tilde{\pi}}_j$	.7961 (.7961)	.8304 (.8305)	.8656 (.8657)	.9036 (.9036)	.9606 (.9606)
μ_j	1.0877 (1.0878)	1.1290 (1.1291)	1.1745 (1.1746)	1.2287 (1.2287)	1.3280 (1.3279)
σ_j	.1033 (.1029)	.1020 (.1018)	.1011 (.1010)	.1007 (.1007)	.1020 (.1018)

These comparisons clearly show that there are practically no differences between the approximate and exact results. Accordingly, the posterior probabilities prob($\pi_j \geq \pi_o | \underset{\sim}{g}$) approximated by our present analysis are not expected to differ significantly from the exact probabilities in our previous analysis. This is so because the normal approximations to these probabilities have been found adequately accurate. The posterior probabilities for π_o = .70(.05).90 computed from formulas (9) and (10) using the current approximate estimates of ρ^* and σ^{*2} are presented below to compare with the exact probabilities (enclosed in parentheses) obtained by numerical integrations with the program MARPRO:

	$x_j = 8$	$x_j = 9$	$x_j = 10$	$x_j = 11$	$x_j = 12$
π_o	$\text{prob}(\pi_j \geq \pi_o \mid \underset{\sim}{g})$				
.90	.0592 (.0557)	.1197 (.1174)	.2306 (.2298)	.4201 (.4193)	.7807 (.7808)
.85	.2042 (.2035)	.3327 (.3348)	.5055 (.5082)	.7096 (.7110)	.9356 (.9379)
.80	.4255 (.4304)	.5850 (.5898)	.7475 (.7502)	.8864 (.8873)	.9848 (.9861)
.75	.6525 (.6588)	.7887 (.7917)	.8960 (.8966)	.9643 (.9642)	.9970 (.9974)
.70	.8249 (.8276)	.9117 (.9115)	.9651 (.9642)	.9908 (.9905)	.9995 (>.999)

The small discrepancies between the exact and approximate probabilities in the above table will not have effects on our decision making in practical applications.

Sometimes, credibility intervals may be of interest to an investigator. They can be approximated using formulas (11) and (12). For our present example, we have computed the approximate posterior 95% credibility intervals of π_j for each observed x_j:

x_j	95% confidence interval of π_j
8	(.5991 .9233)
9	(.6417 .9426)
10	(.6863 .9613)
11	(.7361 .9792)
12	(.8165 .9982)

The reader may check the exact posterior probabilities given in Table 7 of Lewis, Wang, and Novick (1973) to convince himself that these approximate intervals are sufficiently close to the exact intervals which are very difficult to obtain directly from the actual posterior marginal density functions of γ_j .

In conclusion, it is felt that these tables will prove useful in analyses of m-group proportion data (with equal sample size n) without recourse to the program MARPRO.

References

Lewis, C., Wang, M. M., & Novick, M. R. Marginal distributions for the estimation of proportions in m groups. ACT Technical Bulletin No. 13. Iowa City, Iowa: The American College Testing Program, 1973.

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-45.

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 10

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.2456	.2578	.2717	.2877	.3060	.2823	.2997	.3198	.3432	.3702
	.0670	.0742	.0826	.0928	.1051	.0874	.0987	.1126	.1298	.1507
.02	.3642	.3798	.3967	.4150	.4348	.4081	.4282	.4500	.4735	.4984
	.1413	.1537	.1678	.1834	.2011	.1757	.1934	.2133	.2357	.2605
.03	.4425	.4584	.4751	.4927	.5110	.4883	.5077	.5279	.5489	.5703
	.2048	.2197	.2358	.2534	.2720	.2474	.2672	.2885	.3114	.3355
.04	.5001	.5153	.5311	.5473	.5638	.5460	.5639	.5821	.6006	.6191
	.2589	.2746	.2915	.3091	.3277	.3065	.3266	.3476	.3695	.3921
.05	.5450	.5593	.5739	.5887	.6036	.5902	.6065	.6230	.6393	.6555
	.3053	.3213	.3381	.3553	.3731	.3560	.3756	.3959	.4165	.4373
			(N = 12)					(N = 14)		
.01	.3153	.3382	.3650	.3960	.4312	.3455	.3739	.4072	.4453	.4874
	.1078	.1243	.1447	.1701	.2008	.1284	.1504	.1781	.2122	.2526
.02	.4460	.4702	.4964	.5240	.5527	.4791	.5071	.5369	.5677	.5987
	.2082	.2312	.2572	.2859	.3171	.2388	.2672	.2988	.3330	.3690
.03	.5268	.5491	.5721	.5955	.6188	.5597	.5843	.6094	.6344	.6587
	.2862	.3106	.3366	.3640	.3922	.3217	.3500	.3801	.4110	.4421
.04	.5838	.6037	.6238	.6437	.6633	.6156	.6370	.6583	.6790	.6990
	.3487	.3725	.3971	.4222	.4477	.3864	.4132	.4407	.4680	.4953
.05	.6269	.6447	.6622	.6795	.6962	.6574	.6761	.6944	.7120	.7288
	.4001	.4227	.4455	.4685	.4913	.4387	.4635	.4884	.5128	.5367
			(N = 16)					(N = 18)		
.01	.3732	.4071	.4466	.4907	.5376	.3988	.4380	.4829	.5319	.5816
	.1487	.1768	.2122	.2547	.3033	.1688	.2032	.2460	.2964	.3515
.02	.5085	.5397	.5723	.6054	.6377	.5347	.5687	.6035	.6379	.6707
	.2678	.3010	.3375	.3765	.4163	.2949	.3328	.3736	.4160	.4580
.03	.5882	.6147	.6413	.6671	.6918	.6132	.6412	.6687	.6950	.7196
	.3540	.3860	.4193	.4527	.4859	.3836	.4186	.4545	.4899	.5242
.04	.6428	.6653	.6873	.7084	.7283	.6663	.6896	.7120	.7332	.7528
	.4201	.4494	.4789	.5080	.5362	.4504	.4817	.5128	.5431	.5718
.05	.6832	.7025	.7211	.7388	.7554	.7053	.7250	.7437	.7612	.7775
	.4727	.4993	.5255	.5510	.5755	.5028	.5308	.5580	.5840	.6087

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 10

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.4225 .1885	.4667 .2293	.5164 .2793	.5688 .3363	.6197 .3959	.4446 .2078	.4933 .2548	.5470 .3114	.6016 .3738	.6525 .4364
.02	.5583 .3204	.5946 .3624	.6311 .4071	.6662 .4521	.6989 .4961	.5797 .3445	.6178 .3902	.6554 .4377	.6909 .4848	.7230 .5294
.03	.6353 .4108	.6644 .4484	.6925 .4863	.7189 .5230	.7431 .5578	.6550 .4358	.6850 .4757	.7134 .5150	.7396 .5525	.7632 .5875
.04	.6869 .4777	.7107 .5108	.7333 .5430	.7542 .5737	.7734 .6025	.7051 .5027	.7292 .5369	.7517 .5699	.7723 .6008	.7909 .6294
.05	.7245 .5299	.7444 .5588	.7630 .5866	.7802 .6127	.7960 .6373	.7413 .5541	.7612 .5837	.7797 .6118	.7965 .6380	.8118 .6622
			(N = 24)					(N = 26)		
.01	.4653 .2266	.5181 .2798	.5750 .3423	.6308 .4088	.6808 .4729	.4846 .2449	.5412 .3040	.6005 .3717	.6566 .4411	.7052 .5056
.02	.5991 .3670	.6388 .4161	.6771 .4661	.7124 .5143	.7438 .5592	.6169 .3884	.6578 .4403	.6965 .4921	.7314 .5411	.7618 .5856
.03	.6728 .4591	.7033 .5006	.7317 .5410	.7576 .5790	.7805 .6136	.6887 .4803	.7196 .5234	.7480 .5646	.7733 .6025	.7955 .6367
.04	.7212 .5252	.7455 .5606	.7679 .5940	.7880 .6248	.8060 .6531	.7357 .5461	.7600 .5820	.7821 .6156	.8018 .6464	.8192 .6742
.05	.7562 .5760	.7760 .6061	.7942 .6343	.8106 .6603	.8253 .6839	.7694 .5959	.7891 .6262	.8070 .6544	.8229 .6800	.8371 .7032
			(N = 28)					(N = 30)		
.01	.5028 .2628	.5626 .3273	.6236 .3993	.6795 .4708	.7264 .5351	.5198 .2801	.5825 .3497	.6447 .4254	.6998 .4979	.7449 .5615
.02	.6332 .4084	.6750 .4627	.7138 .5130	.7482 .5654	.7776 .6094	.6482 .4273	.6907 .4838	.7295 .5382	.7632 .5876	.7916 .6308
.03	.7032 .5002	.7343 .5444	.7625 .5860	.7873 .6239	.8087 .6575	.7165 .5187	.7476 .5637	.7754 .6055	.7996 .6431	.8204 .6762
.04	.7488 .5652	.7730 .6015	.7947 .6351	.8139 .6656	.8307 .6928	.7606 .5827	.7846 .6193	.8060 .6529	.8247 .6830	.8409 .7096
.05	.7813 .6140	.8008 .6445	.8183 .6725	.8338 .6978	.8474 .7203	.7920 .6306	.8112 .6610	.8283 .6887	.8434 .7136	.8565 .7356

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 15

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.2297 .0582	.2447 .0663	.2625 .0765	.2839 .0898	.3095 .1068	.2642 .0760	.2858 .0892	.3121 .1066	.3440 .1294	.3822 .1592
.02	.3388 .1218	.3578 .1359	.3792 .1527	.4031 .1723	.4294 .1952	.3805 .1523	.4056 .1729	.4336 .1974	.4645 .2260	.4977 .2585
.03	.4116 .1767	.4313 .1940	.4526 .2135	.4754 .2352	.4994 .2590	.4558 .2153	.4803 .2388	.5065 .2651	.5340 .2941	.5623 .3252
.04	.4659 .2244	.4851 .2430	.5054 .2635	.5265 .2856	.5482 .3091	.5108 .2680	.5338 .2924	.5577 .3187	.5821 .3466	.6066 .3757
.05	.5089 .2660	.5272 .2852	.5461 .3057	.5655 .3275	.5852 .3502	.5536 .3132	.5749 .3373	.5967 .3630	.6184 .3893	.6399 .4163
			(N = 12)					(N = 14)		
.01	.2956 .0943	.3243 .1137	.3597 .1398	.4024 .1742	.4516 .2177	.3246 .1128	.3606 .1392	.4051 .1753	.4573 .2218	.5143 .2778
.02	.4170 .1816	.4477 .2090	.4817 .2415	.5182 .2785	.5561 .3193	.4494 .2098	.4851 .2439	.5241 .2839	.5648 .3284	.6052 .3754
.03	.4935 .2509	.5221 .2805	.5522 .3131	.5831 .3482	.6138 .3849	.5261 .2840	.5582 .3191	.5913 .3572	.6243 .3973	.6560 .4374
.04	.5484 .3075	.5744 .3369	.6010 .3663	.6275 .4008	.6533 .4335	.5804 .3434	.6088 .3772	.6372 .4125	.6649 .4483	.6911 .4834
.05	.5906 .3551	.6141 .3834	.6377 .4129	.6608 .4428	.6831 .4724	.6215 .3922	.6469 .4243	.6716 .4566	.6954 .4889	.7178 .5202
			(N = 16)					(N = 18)		
.01	.3514 .1313	.3949 .1656	.4477 .2119	.5075 .2700	.5685 .3352	.3765 .1499	.4271 .1925	.4873 .2490	.5522 .3165	.6141 .3876
.02	.4783 .2366	.5186 .2774	.5616 .3242	.6049 .3745	.6462 .4256	.5045 .2622	.5487 .3093	.5947 .3620	.6394 .4166	.6805 .4701
.03	.5547 .3146	.5897 .3548	.6249 .3975	.6590 .4410	.6909 .4834	.5801 .3432	.6173 .3878	.6539 .4341	.6885 .4800	.7199 .5236
.04	.6081 .3759	.6383 .4135	.6680 .4521	.6961 .4901	.7221 .5264	.6323 .4056	.6639 .4464	.6942 .4872	.7224 .5267	.7478 .5635
.05	.6482 .4256	.6747 .4605	.7001 .4951	.7240 .5288	.7460 .5607	.6713 .4556	.6985 .4927	.7242 .5290	.7479 .5635	.7693 .5955

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 15

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.3999	.4573	.5237	.5915	.6523	.4219	.4855	.5567	.6256	.6842
	.1683	.2193	.2855	.3605	.4345	.1865	.2459	.3205	.4010	.4759
.02	.5282	.5758	.6239	.6691	.7092	.5499	.6002	.6497	.6948	.7336
	.2865	.3394	.3970	.4547	.5090	.3097	.3677	.4292	.4889	.5435
.03	.6027	.6417	.6792	.7137	.7442	.6231	.6634	.7014	.7354	.7649
	.3696	.4181	.4672	.5147	.5584	.3943	.4459	.4974	.5455	.5891
.04	.6537	.6862	.7168	.7447	.7694	.6727	.7059	.7365	.7639	.7877
	.4327	.4761	.5186	.5589	.5958	.4576	.5031	.5467	.5873	.6238
.05	.6915	.7192	.7449	.7682	.7889	.7093	.7373	.7628	.7856	.8055
	.4828	.5216	.5589	.5937	.6256	.5074	.5476	.5855	.6204	.6516
			(N = 24)					(N = 26)		
.01	.4427	.5119	.5866	.6552	.7111	.4622	.5365	.6135	.6810	.7339
	.2046	.2720	.3541	.4378	.5124	.2223	.2976	.3857	.4714	.5444
.02	.5697	.6223	.6726	.7171	.7544	.5880	.6423	.6930	.7366	.7724
	.3317	.3944	.4589	.5198	.5737	.3526	.4192	.4862	.5476	.6006
.03	.6414	.6828	.7208	.7542	.7827	.6582	.7002	.7381	.7707	.7980
	.4171	.4716	.5245	.5730	.6161	.4387	.4954	.5492	.5978	.6399
.04	.6897	.7234	.7538	.7805	.8035	.7051	.7389	.7690	.7950	.8171
	.4804	.5277	.5721	.6126	.6485	.5017	.5500	.5949	.6350	.6702
.05	.7252	.7532	.7785	.8006	.8198	.7354	.7674	.7922	.8137	.8321
	.5299	.5709	.6093	.6438	.6745	.5504	.5922	.6305	.6646	.6945
			(N = 28)					(N = 30)		
.01	.4807	.5594	.6377	.7034	.7535	.4981	.5806	.6594	.7231	.7705
	.2397	.3224	.4153	.5015	.5728	.2567	.3462	.4427	.5289	.5981
.02	.6048	.6606	.7112	.7537	.7879	.6204	.6772	.7276	.7689	.8016
	.3724	.4427	.5113	.5726	.6244	.3913	.4645	.5344	.5952	.6458
.03	.6734	.7159	.7534	.7852	.8114	.6874	.7301	.7672	.7980	.8231
	.4587	.5172	.5717	.6199	.6612	.4774	.5374	.5923	.6398	.6800
.04	.7190	.7529	.7826	.8078	.8290	.7317	.7655	.7946	.8191	.8394
	.5212	.5706	.6157	.6552	.6895	.5393	.5894	.6343	.6733	.7066
.05	.7522	.7801	.8044	.8253	.8430	.7639	.7915	.8153	.8355	.8525
	.5593	.6117	.6497	.6833	.7125	.5867	.6293	.6671	.7001	.7285

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 20

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.2173	.2342	.2550	.2809	.3131	.2500	.2746	.3058	.3452	.3933
	.0517	.0603	.0718	.0874	.1085	.0677	.0820	.1018	.1296	.1671
.02	.3191	.3406	.3654	.3937	.4255	.3591	.3877	.42[illegible]7	.4577	.4978
	.1077	.1228	.1413	.1639	.1909	.1353	.1576	.1854	.2187	.2576
.03	.3877	.4101	.4349	.4618	.4907	.4305	.4588	.4897	.5226	.5565
	.1565	.1751	.1967	.2215	.2495	.1917	.2176	.2474	.2811	.3178
.04	.4394	.4614	.4851	.5102	.5362	.4833	.5102	.5385	.5678	.5972
	.1993	.2196	.2425	.2678	.2952	.2398	.2669	.2969	.3294	.3635
.05	.4806	.5019	.5242	.5474	.5710	.5248	.5500	.5760	.6022	.6282
	.2371	.2583	.2815	.3065	.3330	.2813	.3086	.3380	.3688	.4007
			(N = 12)					(N = 14)		
.01	.2801	.3132	.3557	.4085	.4693	.3081	.3500	.4038	.4680	.5359
	.0843	.1056	.1360	.1784	.2328	.1012	.1307	.1734	.2306	.2987
.02	.3945	.4299	.4702	.5142	.5594	.4261	.4679	.5144	.5632	.6106
	.1622	.1924	.2296	.2734	.3219	.1883	.2266	.2729	.3256	.3807
.03	.4674	.5009	.5368	.5738	.6104	.4997	.5377	.5774	.6170	.6545
	.2249	.2579	.2955	.3366	.3798	.2560	.2958	.3403	.3874	.4346
.04	.5205	.5513	.5832	.6152	.6461	.5525	.5865	.6209	.6543	.6856
	.2769	.3102	.3465	.3848	.4234	.3111	.3499	.3914	.4337	.4752
.05	.5616	.5899	.6185	.6465	.6734	.5930	.6237	.6539	.6828	.7097
	.3210	.3537	.3882	.4235	.4587	.3569	.3943	.4327	.4710	.5081
			(N = 16)					(N = 18)		
.01	.3343	.3852	.4493	.5215	.5913	.3589	.4185	.4915	.5683	.6366
	.1185	.1572	.2125	.2830	.3595	.1358	.1842	.2520	.3330	.4136
.02	.4547	.5022	.5536	.6052	.6529	.4807	.5332	.5883	.6411	.6877
	.2135	.2598	.3144	.3739	.4331	.2378	.2917	.3535	.4178	.4787
.03	.5283	.5700	.6125	.6533	.6906	.5539	.5986	.6429	.6840	.7204
	.2852	.3313	.3816	.4328	.4822	.3127	.3644	.4191	.4732	.5236
.04	.5804	.6170	.6530	.6870	.7179	.6051	.6436	.6806	.7145	.7446
	.3424	.3862	.4317	.4769	.5198	.3713	.4193	.4680	.5148	.5582
.05	.6201	.6525	.6837	.7128	.7392	.6438	.6774	.7091	.7380	.7636
	.3894	.4306	.4720	.5123	.5502	.4191	.4633	.5069	.5483	.5863

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 20

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.3821	.4501	.5301	.6084	.6736	.4041	.4798	.5648	.6427	.7041
	.1533	.2118	.2910	.3791	.4607	.1708	.2394	.3283	.4209	.5017
.02	.5045	.5612	.6188	.6717	.7167	.5264	.5866	.6458	.6981	.7410
	.2611	.3220	.3897	.4572	.5187	.2836	.3509	.4234	.4926	.5534
.03	.5768	.6240	.6694	.7102	.7454	.5976	.6466	.6925	.7327	.7664
	.3384	.3951	.4534	.5090	.5595	.3626	.4234	.4844	.5409	.5908
.04	.6270	.6670	.7044	.7379	.7668	.6466	.6876	.7251	.7579	.7857
	.3980	.4497	.5005	.5483	.5913	.4228	.4772	.5297	.5778	.6201
.05	.6647	.6992	.7310	.7593	.7840	.6833	.7183	.7499	.7776	.8012
	.4461	.4930	.5381	.5797	.6175	.4709	.5197	.5657	.6075	.6443
			(N = 24)					(N = 26)		
.01	.4249	.5075	.5959	.6720	.7295	.4446	.5334	.6236	.6971	.7510
	.1881	.2667	.3637	.4584	.5372	.2053	.2933	.3968	.4919	.5683
.02	.5465	.6097	.6697	.7208	.7617	.5651	.6307	.6909	.7405	.7794
	.3050	.3781	.4542	.5243	.5839	.3255	.4039	.4825	.5525	.6107
.03	.6165	.6669	.7129	.7522	.7844	.6337	.6852	.7310	.7691	.7998
	.3853	.4498	.5126	.5694	.6183	.4066	.4741	.5384	.5947	.6423
.04	.6643	.7060	.7433	.7752	.8018	.6803	.7224	.7593	.7903	.8158
	.4457	.5025	.5560	.6039	.6454	.4670	.5256	.5797	.6273	.6677
.05	.6999	.7352	.7665	.7933	.8160	.7149	.7502	.7810	.8071	.8288
	.4937	.5439	.5905	.6318	.6680	.5146	.5659	.6127	.6536	.6888
			(N = 28)					(N = 30)		
.01	.4633	.5574	.6483	.7189	.7694	.4811	.5797	.6702	.7378	.7852
	.2223	.3192	.4275	.5220	.5957	.2391	.3442	.4557	.5489	.6197
.02	.5824	.6498	.7097	.7578	.7947	.5985	.6673	.7266	.7730	.8080
	.3452	.4279	.5085	.5780	.6343	.3640	.4506	.5322	.6008	.6554
.03	.6495	.7017	.7470	.7839	.8133	.6641	.7167	.7613	.7971	.8251
	.4267	.4967	.5616	.6174	.6638	.4456	.5177	.5829	.6380	.6828
.04	.6949	.7372	.7736	.8036	.8279	.7082	.7506	.7863	.8153	.8386
	.4868	.5470	.6014	.6482	.6873	.5052	.5666	.6209	.6668	.7049
.05	.7284	.7637	.7940	.8192	.8399	.7407	.7759	.8055	.8299	.8498
	.5339	.5861	.6328	.6731	.7070	.5517	.6047	.6510	.6905	.7237

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 25

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.2070 .0466	.2253 .0556	.2486 .0680	.2784 .0855	.3167 .1104	.2384 .0613	.2653 .0762	.3006 .0981	.3465 .1300	.4035 .1745
.02	.3031 .0969	.3264 .1125	.3539 .1322	.3859 .1570	.4225 .1876	.3417 .1223	.3730 .1456	.4100 .1757	.4523 .2131	.4983 .2572
.03	.3683 .1410	.3928 .1604	.4203 .1835	.4508 .2107	.4836 .2419	.4099 .1736	.4412 .2010	.4760 .2335	.5134 .2709	.5521 .3122
.04	.4178 .1800	.4421 .2015	.4686 .2261	.4969 .2538	.5266 .2844	.4608 .2178	.4907 .2467	.5228 .2796	.5563 .3159	.5899 .3543
.05	.4575 .2147	.4811 .2373	.5063 .2624	.5326 .2900	.5596 .3196	.5010 .2562	.5294 .2858	.5590 .3182	.5891 .3527	.6189 .3886
			(N = 12)					(N = 14)		
.01	.2674 .0765	.3038 .0990	.3524 .1331	.4142 .1824	.4844 .2460	.2946 .0923	.3412 .1239	.4030 .1720	.4773 .2384	.5529 .3156
.02	.3761 .1472	.4153 .1793	.4609 .2202	.5112 .2695	.5625 .3245	.4071 .1716	.4537 .2127	.5067 .2643	.5622 .3237	.6152 .3854
.03	.4461 .2046	.4835 .2401	.5243 .2816	.5666 .3278	.6081 .3763	.4781 .2342	.5208 .2773	.5662 .3269	.6114 .3799	.6536 .4327
.04	.4975 .2529	.5323 .2890	.5687 .3293	.6054 .3723	.6406 .4159	.5295 .2856	.5682 .3283	.6077 .3747	.6461 .4225	.6816 .4692
.05	.5378 .2942	.5699 .3300	.6027 .3684	.6351 .4085	.6659 .4482	.5693 .3289	.6045 .3703	.6394 .4135	.6728 .4571	.7036 .4991
			(N = 16)					(N = 18)		
.01	.3202 .1084	.3772 .1503	.4511 .2132	.5331 .2940	.6082 .3782	.3444 .1248	.4116 .1777	.4955 .2550	.5809 .3461	.6525 .4326
.02	.4354 .1956	.4887 .2457	.5474 .3068	.6057 .3737	.6582 .4391	.4613 .2188	.5205 .2776	.5834 .3471	.6426 .4189	.6933 .4856
.03	.5066 .2621	.5539 .3127	.6026 .3689	.6490 .4266	.6906 .4816	.5322 .2886	.5833 .3458	.6341 .4074	.6808 .4682	.7211 .5240
.04	.5575 .3158	.5995 .3645	.6411 .4159	.6800 .4669	.7148 .5148	.5825 .3441	.6269 .3978	.6697 .4529	.7085 .5059	.7422 .5543
.05	.5967 .3606	.6342 .4067	.6705 .4539	.7041 .4997	.7341 .5423	.6208 .3897	.6600 .4397	.6970 .4896	.7303 .5367	.7593 .5794

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 25

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.3674	.4444	.5357	.6212	.6882	.3893	.4753	.5717	.6551	.7174
	.1414	.2060	.2960	.3934	.4792	.1582	.2344	.3351	.4357	.5194
.02	.4851	.5495	.6151	.6740	.7222	.5071	.5758	.6430	.7007	.7464
	.2413	.3085	.3845	.4595	.5258	.2631	.3377	.4190	.4956	.5607
.03	.5554	.6095	.6617	.7078	.7464	.5765	.6330	.6857	.7309	.7676
	.3137	.3768	.4426	.5051	.5605	.3374	.4056	.4746	.5378	.5921
.04	.6048	.6512	.6945	.7327	.7651	.6249	.6727	.7162	.7534	.7843
	.3704	.4285	.4863	.5402	.5883	.3948	.4566	.5165	.5706	.6176
.05	.6422	.6826	.7198	.7525	.7803	.6614	.7026	.7396	.7714	.7981
	.4164	.4697	.5215	.5693	.6114	.4412	.4971	.5501	.5977	.6392
			(N = 24)					(N = 26)		
.01	.4102	.5043	.6035	.6839	.7416	.4301	.5313	.6315	.7084	.7621
	.1751	.2626	.3717	.4733	.5540	.1919	.2903	.4056	.5066	.5842
.02	.5274	.5997	.6677	.7236	.7669	.5464	.6215	.6894	.7435	.7843
	.2840	.3655	.4509	.5276	.5912	.3042	.3918	.4799	.5563	.6178
.03	.5958	.6542	.7069	.7508	.7857	.6134	.6732	.7256	.7680	.8013
	.3598	.4326	.5037	.5669	.6198	.3809	.4575	.5301	.5926	.6443
.04	.6431	.6918	.7351	.7713	.8008	.6596	.7090	.7518	.7869	.8149
	.4177	.4824	.5437	.5976	.6435	.4390	.5062	.5681	.6216	.6660
.05	.6786	.7203	.7569	.7878	.8132	.6941	.7361	.7722	.8021	.8263
	.4641	.5220	.5757	.6229	.6632	.4851	.5448	.5988	.6454	.6844
			(N = 28)					(N = 30)		
.01	.4491	.5563	.6563	.7294	.7796	.4672	.5794	.6783	.7477	.7946
	.2086	.3172	.4368	.5362	.6107	.2252	.3430	.4656	.5628	.6339
.02	.5640	.6414	.7088	.7608	.7994	.5804	.6595	.7260	.7760	.8125
	.3236	.4166	.5066	.5819	.6413	.3422	.4398	.5309	.6050	.6622
.03	.6297	.6904	.7422	.7832	.8147	.6447	.7061	.7570	.7965	.8266
	.4009	.4807	.5541	.6159	.6656	.4198	.5023	.5760	.6367	.6850
.04	.6747	.7245	.7666	.8005	.8273	.6886	.7385	.7798	.8125	.8381
	.4589	.5281	.5904	.6429	.6861	.4777	.5484	.6105	.6621	.7039
.05	.7083	.7503	.7857	.8146	.8378	.7212	.7631	.7978	.8256	.8478
	.5048	.5657	.6195	.6654	.7034	.5231	.5848	.6385	.6832	.7201

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 30

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.1984	.2178	.2430	.2764	.3202	.2286	.2572	.2961	.3480	.4128
	.0427	.0517	.0647	.0839	.1122	.0562	.0715	.0948	.1305	.1814
.02	.2897	.3144	.3441	.3794	.4201	.3271	.3606	.4010	.4479	.4990
	.0883	.1041	.1248	.1514	.1850	.1118	.1359	.1678	.2085	.2572
.03	.3521	.3781	.4079	.4414	.4779	.3927	.4263	.4644	.5058	.5487
	.1287	.1485	.1726	.2017	.2358	.1592	.1874	.2221	.2626	.3079
.04	.3997	.4257	.4545	.4858	.5187	.4418	.4743	.5096	.5467	.5839
	.1646	.1866	.2125	.2423	.2755	.2001	.2304	.2655	.3048	.3467
.05	.4381	.4636	.4911	.5202	.5501	.4810	.5119	.5447	.5782	.6113
	.1967	.2201	.2468	.2764	.3085	.2361	.2671	.3020	.3396	.3789
			(N = 12)					(N = 14)		
.01	.2567	.2958	.3498	.4195	.4971	.2832	.3336	.4027	.4854	.5663
	.0703	.0936	.1307	.1861	.2574	.0850	.1181	.1712	.2453	.3293
.02	.3607	.4029	.4532	.5090	.5652	.3911	.4417	.5003	.5617	.6191
	.1352	.1685	.2126	.2667	.3268	.1583	.2014	.2573	.3224	.3895
.03	.4282	.4688	.5138	.5607	.6064	.4599	.5066	.5569	.6069	.6531
	.1885	.2255	.2702	.3206	.3736	.2166	.2622	.3159	.3739	.4315
.04	.4781	.5162	.5566	.5973	.6363	.5099	.5527	.5967	.6394	.6784
	.2334	.2718	.3153	.3622	.4099	.2647	.3105	.3611	.4135	.4644
.05	.5175	.5529	.5895	.6256	.6598	.5491	.5881	.6273	.6646	.6987
	.2724	.3105	.3524	.3961	.4397	.3059	.3504	.3979	.4458	.4919
			(N = 16)					(N = 18)		
.01	.3083	.3704	.4529	.5427	.6209	.3322	.4058	.4991	.5909	.6640
	.1002	.1446	.2142	.3032	.3925	.1159	.1724	.2578	.3567	.4466
.02	.4191	.4773	.5424	.6064	.6624	.4448	.5099	.5796	.6441	.6976
	.1810	.2341	.3009	.3738	.4439	.2032	.2662	.3421	.4203	.4909
.03	.4882	.5403	.5944	.6457	.6908	.5139	.5704	.6270	.6784	.7217
	.2433	.2973	.3587	.4218	.4814	.2690	.3306	.3980	.4645	.5244
.04	.5381	.5847	.6311	.6744	.7125	.5632	.6129	.6608	.7038	.7405
	.2942	.3467	.4029	.4589	.5113	.3216	.3801	.4408	.4989	.5513
.05	.5767	.6186	.6594	.6970	.7302	.6012	.6452	.6869	.7241	.7560
	.3368	.3869	.4388	.4894	.5363	.3654	.4202	.4754	.5274	.5741

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 30

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.3550 .1318	.4397 .2012	.5406 .3003	.6310 .4046	.6986 .4926	.3769 .1481	.4718 .2304	.5774 .3408	.6645 .4471	.7267 .5319
.02	.4686 .2250	.5397 .2973	.6123 .3805	.6759 .4614	.7265 .5314	.4907 .2462	.5668 .3270	.6410 .4160	.7029 .4981	.7505 .5664
.03	.5372 .2934	.5974 .3618	.6555 .4342	.7060 .5021	.7472 .5613	.5586 .3166	.6217 .3911	.6803 .4668	.7295 .5355	.7686 .5933
.04	.5858 .3474	.6378 .4110	.6865 .4750	.7287 .5341	.7638 .5860	.6063 .3717	.6601 .4396	.7088 .5057	.7500 .5652	.7833 .6158
.05	.6230 .3919	.6686 .4506	.7106 .5082	.7470 .5607	.7775 .6068	.6425 .4164	.6893 .4784	.7311 .5374	.7665 .5899	.7957 .6350
			(N = 24)					(N = 26)		
.01	.3978 .1644	.5018 .2595	.6096 .3782	.6926 .4844	.7501 .5659	.4179 .1809	.5297 .2880	.6378 .4127	.7165 .5174	.7698 .5954
.02	.5113 .2667	.5915 .3553	.6662 .4484	.7260 .5306	.7707 .5966	.5305 .2866	.6140 .3821	.6885 .4782	.7459 .5595	.7880 .6232
.03	.5782 .3388	.6435 .4184	.7021 .4965	.7498 .5651	.7868 .6213	.5962 .3598	.6632 .4438	.7214 .5237	.7673 .5912	.8024 .6457
.04	.6248 .3943	.6799 .4659	.7284 .5336	.7683 .5927	.8000 .6419	.6418 .4156	.6978 .4902	.7457 .5588	.7842 .6171	.8143 .6648
.05	.6602 .4393	.7077 .5038	.7491 .5638	.7834 .6158	.8112 .6597	.6762 .4604	.7242 .5273	.7650 .5875	.7981 .6389	.8245 .6813
			(N = 28)					(N = 30)		
.01	.4371 .1974	.5555 .3156	.6626 .4443	.7370 .5467	.7866 .6211	.4555 .2139	.5793 .3422	.6844 .4731	.7548 .5728	.8012 .6440
.02	.5484 .3058	.6345 .4074	.7082 .5053	.7632 .5852	.8028 .6465	.5652 .3244	.6532 .4312	.7257 .5300	.7783 .6082	.8158 .6672
.03	.6128 .3797	.6811 .4676	.7384 .5482	.7826 .6147	.8159 .6674	.6282 .3986	.6973 .4897	.7536 .5706	.7961 .6358	.8277 .6866
.04	.6573 .4355	.7139 .5128	.7610 .5816	.7981 .6389	.8268 .6851	.6716 .4543	.7284 .5334	.7747 .6024	.8104 .6584	.8377 .7030
.05	.6908 .4802	.7389 .5486	.7790 .6089	.8110 .6594	.8361 .7004	.7042 .4987	.7523 .5684	.7915 .6284	.8223 .6777	.8463 .7174

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 40

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.1845	.2053	.2337	.2731	.3270	.2128	.2440	.2886	.3511	.4287
	.0366	.0456	.0595	.0814	.1159	.0484	.0639	.0896	.1318	.1935
.02	.2683	.2949	.3280	.3688	.4167	.3037	.3403	.3863	.4411	.5006
	.0755	.0914	.1130	.1426	.1811	.0961	.1207	.1552	.2015	.2577
.03	.3261	.3544	.3879	.4264	.4689	.3649	.4021	.4456	.4938	.5437
	.1101	.1301	.1558	.1878	.2264	.1372	.1665	.2041	.2497	.3015
.04	.3706	.3992	.4317	.4677	.5061	.4113	.4476	.4882	.5314	.5749
	.1412	.1639	.1915	.2242	.2618	.1732	.2049	.2433	.2876	.3356
.05	.4069	.4351	.4664	.5000	.5351	.4486	.4835	.5214	.5607	.5995
	.1695	.1937	.2223	.2551	.2915	.2051	.2382	.2765	.3191	.3639
			(N = 12)					(N = 14)		
.01	.2394	.2826	.3457	.4289	.5168	.2647	.3212	.4026	.4987	.5856
	.0608	.0851	.1269	.1929	.2754	.0740	.1091	.1701	.2569	.3495
.02	.3358	.3828	.4408	.5058	.5699	.3654	.4223	.4904	.5613	.6250
	.1169	.1518	.2006	.2624	.3309	.1378	.1838	.2467	.3210	.3956
.03	.3993	.4449	.4970	.5517	.6042	.4303	.4835	.5422	.6003	.6526
	.1636	.2028	.2524	.3098	.3701	.1894	.2387	.2991	.3652	.4301
.04	.4468	.4900	.5370	.5848	.6300	.4782	.5274	.5791	.6291	.6739
	.2037	.2447	.2932	.3467	.4012	.2327	.2826	.3398	.3999	.4577
.05	.4845	.5253	.5681	.6107	.6507	.5160	.5615	.6079	.6519	.6914
	.2386	.2801	.3270	.3772	.4272	.2701	.3193	.3734	.4286	.4812
			(N = 16)					(N = 18)		
.01	.2890	.3594	.4564	.5574	.6383	.3123	.3966	.5054	.6055	.6794
	.0878	.1357	.2162	.3177	.4126	.1021	.1641	.2628	.3725	.4658
.02	.3927	.4590	.5348	.6078	.6688	.4181	.4929	.5740	.6465	.7039
	.1587	.2162	.2918	.3745	.4514	.1794	.2484	.3348	.4224	.4989
.03	.4584	.5183	.5817	.6410	.6913	.4841	.5497	.6161	.6749	.7227
	.2143	.2734	.3431	.4151	.4813	.2386	.3068	.3839	.4591	.5252
.04	.5063	.5606	.6155	.6660	.7093	.5317	.5900	.6468	.6968	.7382
	.2603	.3185	.3828	.4472	.5061	.2866	.3521	.4220	.4886	.5474
.05	.5439	.5932	.6418	.6861	.7243	.5688	.6211	.6709	.7146	.7512
	.2995	.3557	.4155	.4738	.5273	.3270	.3893	.4533	.5134	.5665

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 40

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.3348	.4324	.5486	.6448	.7122	.3566	.4664	.5862	.6773	.7389
	.1169	.1939	.3078	.4207	.5106	.1323	.2244	.3496	.4628	.5488
.02	.4419	.5241	.6082	.6789	.7325	.4641	.5526	.6382	.7062	.7562
	.1999	.2800	.3747	.4647	.5394	.2200	.3104	.4116	.5019	.5742
.03	.5076	.5780	.6461	.7035	.7486	.5293	.6036	.6722	.7277	.7702
	.2618	.3384	.4213	.4980	.5628	.2843	.3684	.4554	.5322	.5952
.04	.5547	.6163	.6740	.7228	.7620	.5757	.6398	.6976	.7449	.7820
	.3114	.3835	.4576	.5251	.5827	.3350	.4128	.4895	.5572	.6133
.05	.5911	.6458	.6961	.7387	.7735	.6114	.6678	.7179	.7593	.7922
	.3528	.4203	.4874	.5481	.6002	.3770	.4490	.5179	.5786	.6292
			(N = 24)					(N = 26)		
.01	.3776	.4982	.6188	.7044	.7611	.3979	.5277	.6471	.7274	.7798
	.1478	.2549	.3882	.4997	.5817	.1637	.2849	.4234	.5321	.6101
.02	.4850	.5786	.6643	.7294	.7760	.5046	.6024	.6873	.7493	.7930
	.2398	.3396	.4452	.5348	.6042	.2591	.3674	.4759	.5639	.6305
.03	.5493	.6266	.6950	.7485	.7884	.5679	.6475	.7151	.7664	.8040
	.3056	.3964	.4861	.5627	.6234	.3263	.4229	.5142	.5895	.6479
.04	.5949	.6609	.7183	.7639	.7990	.6125	.679[illegible]	.7365	.7804	.8136
	.3574	.4400	.5186	.5855	.6400	.3785	.4651	.5448	.6108	.6633
.05	.6297	.6873	.7370	.7769	.8082	.6465	.7049	.7538	.7923	.8219
	.3996	.4752	.5455	.6054	.6546	.4209	.4995	.5703	.6293	.6767
			(N = 28)					(N = 30)		
.01	.4175	.5549	.6717	.7471	.7957	.4364	.5797	.6932	.7641	.8096
	.1797	.3139	.4553	.5608	.6348	.1959	.3417	.4842	.5861	.6570
.02	.5230	.6240	.7075	.7665	.8075	.5403	.6437	.7254	.7816	.8202
	.2779	.3936	.5037	.5896	.6536	.2962	.4182	.5290	.6128	.6740
.03	.5851	.6664	.7329	.7820	.8175	.6012	.6836	.7487	.7957	.8293
	.3460	.4474	.5396	.6133	.6696	.3650	.4704	.5629	.6347	.6889
.04	.6287	.6969	.7526	.7948	.8262	.6437	.7125	.7670	.8074	.8372
	.3985	.4885	.5685	.6333	.6838	.4173	.5102	.5902	.6533	.7020
.05	.6619	.7207	.7687	.8057	.8338	.6761	.7350	.7820	.8175	.8443
	.4409	.5218	.5928	.6506	.6963	.4597	.5424	.6132	.6696	.7138

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 50

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.1735	.1953	.2261	.2707	.3332	.2003	.2333	.2827	.3541	.4416
	.0323	.0411	.0554	.0796	.1194	.0427	.0582	.0856	.1332	.2035
.02	.2517	.2795	.3153	.3604	.4144	.2855	.3243	.3748	.4361	.5023
	.0663	.0818	.1041	.1358	.1784	.0847	.1094	.1459	.1964	.2586
.03	.3059	.3358	.3719	.4146	.4622	.3432	.3830	.4308	.4847	.5402
	.0968	.1167	.1429	.1773	.2194	.1212	.1509	.1905	.2402	.2970
.04	.3480	.3784	.4137	.4536	.4965	.3874	.4265	.4713	.5197	.5683
	.1284	.1471	.1755	.2107	.2516	.1535	.1859	.2266	.2748	.3275
.05	.3825	.4127	.4468	.4843	.5235	.4231	.4611	.5031	.5472	.5907
	.1496	.1741	.2038	.2390	.2788	.1824	.2164	.2572	.3036	.3529
			(N = 12)					(N = 14)		
.01	.2258	.2720	.3428	.4367	.5309	.2502	.3113	.4030	.5089	.5983
	.0540	.0786	.1243	.1986	.2888	.0659	.1022	.1696	.2660	.3633
.02	.3164	.3668	.4312	.5038	.5736	.3451	.4069	.4830	.5614	.6294
	.1036	.1391	.1915	.2596	.3342	.1228	.1704	.2388	.3203	.4003
.03	.3767	.4260	.4839	.5451	.6029	.4070	.4653	.5309	.5957	.6524
	.1455	.1858	.2391	.3019	.3678	.1692	.2208	.2865	.3591	.4292
.04	.4221	.4693	.5217	.5753	.6255	.4531	.5074	.5656	.6216	.6708
	.1817	.2262	.2765	.3352	.3951	.2088	.2614	.3239	.3901	.4531
.05	.4585	.5033	.5514	.5994	.6440	.4897	.5404	.5927	.6424	.6863
	.2136	.2571	.3079	.3630	.4181	.2431	.2956	.3549	.4159	.4738
			(N = 16)					(N = 18)		
.01	.2737	.3507	.4597	.5679	.6494	.2966	.3895	.5105	.6154	.6891
	.0785	.1288	.2183	.3283	.4258	.0919	.1578	.2671	.3835	.4782
.02	.3719	.4446	.5293	.6091	.6732	.3971	.4797	.5701	.6485	.7081
	.1421	.2026	.2854	.3754	.4566	.1616	.2350	.3297	.4243	.5042
.03	.4348	.5011	.5722	.6377	.6918	.4604	.5335	.6080	.6726	.7236
	.1928	.2554	.3316	.4103	.4815	.2157	.2887	.3735	.4555	.5260
.04	.4811	.5416	.6035	.6599	.7072	.5066	.5721	.6363	.6918	.7366
	.2350	.2971	.3678	.4387	.5027	.2600	.3309	.4082	.4813	.5447
.05	.5177	.5731	.6282	.6781	.7202	.5428	.6021	.6587	.7078	.7479
	.2713	.3318	.3978	.4626	.5210	.2978	.3657	.4368	.5034	.5613

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 50

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.3188 .1058	.4270 .1805	.5547 .3134	.6539 .4316	.7207 .5221	.3405 .1203	.4626 .2201	.5927 .3563	.6855 .4732	.7464 .5593
.02	.4208 .1811	.5120 .2668	.6055 .3708	.6812 .4672	.7365 .5447	.4431 .2003	.5418 .2980	.6364 .4087	.7086 .5048	.7599 .5793
.03	.4841 .2381	.5629 .3208	.6392 .4120	.7019 .4954	.7496 .5639	.5060 .2596	.5896 .3513	.6663 .4471	.7266 .5302	.7713 .5966
.04	.5299 .2841	.5994 .3627	.6646 .4446	.7187 .5188	.7609 .5808	.5512 .3070	.6240 .3925	.6893 .4777	.7415 .5518	.7812 .6118
.05	.5655 .3228	.6278 .3971	.6851 .4720	.7328 .5391	.7707 .5956	.5862 .3465	.6508 .4253	.7080 .5036	.7541 .5705	.7899 .6253
			(N = 24)					(N = 26)		
.01	.3616 .1354	.4959 .2518	.6253 .3954	.7119 .5096	.7678 .5914	.3821 .1507	.5266 .2829	.6534 .4307	.7343 .5416	.7859 .6192
.02	.4642 .2195	.5689 .3279	.6632 .4431	.7318 .5378	.7795 .6092	.4841 .2384	.5937 .3565	.6866 .4744	.7517 .5671	.7962 .6353
.03	.5264 .2806	.6136 .3799	.6900 .4788	.7477 .5611	.7896 .6250	.5454 .3009	.6355 .4071	.7108 .5076	.7658 .5881	.8052 .6496
.04	.5708 .3289	.6460 .4202	.7109 .5077	.7610 .5809	.7984 .6387	.5889 .3498	.6659 .4461	.7298 .5347	.7778 .6065	.8131 .6622
.05	.6052 .3691	.6714 .4533	.7280 .5321	.7723 .5980	.8062 .6512	.6225 .3902	.6899 .4784	.7456 .5577	.7882 .6227	.8202 .6737
			(N = 28)					(N = 30)		
.01	.4021 .1665	.5547 .3129	.6778 .4628	.7534 .5697	.8013 .6435	.4214 .1825	.5803 .3416	.6990 .4916	.7700 .5947	.8147 .6649
.02	.5029 .2569	.6163 .3835	.7072 .5027	.7689 .5930	.8106 .6583	.5207 .2750	.6368 .4090	.7253 .5285	.7839 .6160	.8231 .6786
.03	.5631 .3204	.6553 .4324	.7291 .5338	.7816 .6124	.8187 .6714	.5796 .3391	.6732 .4560	.7453 .5575	.7955 .6342	.8304 .6906
.04	.6057 .3698	.6839 .4702	.7466 .5593	.7925 .6295	.8258 .6829	.6213 .3888	.7002 .4926	.7615 .5816	.8054 .6499	.8370 .7015
.05	.6385 .4102	.7065 .5013	.7611 .5810	.8020 .6444	.8323 .6936	.6532 .4291	.7216 .5227	.7750 .6021	.8141 .6639	.8429 .7113

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR N = 60

MEAN SQUARED DEVIATIONS OF G

LAMBDA/RHO	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.1646 .0289	.1870 .0376	.2196 .0521	.2687 .0780	.3389 .1227	.1902 .0384	.2244 .0538	.2778 .0824	.3571 .1347	.4519 .2116
.02	.2382 .0591	.2668 .0744	.3047 .0970	.3536 .1304	.4128 .1765	.2706 .0759	.3111 .1005	.3652 .1383	.4323 .1926	.5039 .2595
.03	.2896 .0866	.3205 .1061	.3588 .1328	.4049 .1688	.4569 .2181	.3256 .1089	.3673 .1386	.4187 .1797	.4774 .2327	.5377 .2938
.04	.3297 .1115	.3613 .1339	.3989 .1631	.4421 .2000	.4889 .2437	.3678 .1383	.4091 .1710	.4575 .2134	.5103 .2647	.5632 .3213
.05	.3626 .1344	.3943 .1588	.4308 .1894	.4713 .2262	.5142 .2686	.4022 .1648	.4425 .1992	.4880 .2419	.5363 .2915	.5839 .3446
			(N = 12)					(N = 14)		
.01	.2146 .0486	.2632 .0734	.3405 .1221	.4433 .2036	.5412 .2987	.2382 .0595	.3030 .0965	.4036 .1695	.5167 .2730	.6073 .3732
.02	.3006 .0934	.3536 .1291	.4234 .1844	.5025 .2578	.5766 .3370	.3286 .1112	.3943 .1599	.4771 .2326	.5616 .3200	.6327 .4039
.03	.3582 .1314	.4105 .1724	.4732 .2283	.5400 .2960	.6020 .3662	.3880 .1537	.4503 .2067	.5219 .2766	.5922 .3545	.6524 .4287
.04	.4018 .1645	.4522 .2081	..5093 .2633	.5679 .3263	.6222 .3905	.4325 .1902	.4910 .2447	.5546 .3113	.6158 .3825	.6686 .4498
.05	.4370 .1940	.4851 .2387	.5377 ..2927	.5904 .3520	.6389 .4112	.4681 .2221	.5229 .2767	.5805 .3403	.6350 .4062	.6824 .4682
			(N = 16)					(N = 18)		
.01	.2612 .0713	.3435 .1233	.4626 .2203	.5756 .3362	.6570 .4350	.2837 .0839	.3837 .1528	.5147 .2706	.6224 .3914	.6956 .4866
.02	.3549 .1294	.4328 .1918	.5251 .2804	.6102 .3761	.6765 .4696	.3798 .1477	.4689 .2243	.5672 .3259	.6500 .4257	.7112 .5081
.03	.4155 .1759	.4869 .2410	.5646 .3226	.6353 .4068	.6923 .4818	.4409 .1977	.5203 .2745	.6018 .3657	.6710 .4530	.7243 .5267
.04	.4603 .2150	.5259 .2801	.5980 .3561	.6553 .4322	.7056 .5002	.4858 .2391	.5574 .3140	.6279 .3973	.6881 .4759	.7355 .5429
.05	.4960 .2489	.5565 .3128	.6173 .3841	.6719 .4539	.7172 .5165	.5212 .2745	.5864 .3468	.6490 .4238	.7026 .4958	.7454 .5573

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR N = 60

MEAN SQUARED DEVIATIONS OF G

LAMBDA/MU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.3057 .0972	.4220 .1645	.5595 .3180	.6602 .4392	.7264 .5299	.3273 .1110	.4598 .2170	.5976 .3614	.6913 .4606	.7515 .5666
.02	.4034 .1663	.5024 .2567	.6036 .3680	.6829 .4691	.7396 .5486	.4258 .1849	.5332 .2884	.6351 .4067	.7184 .5070	.7626 .5832
.03	.4646 .2192	.5507 .3069	.6340 .4051	.7008 .4934	.7504 .5648	.4867 .2402	.5783 .3378	.6619 .4408	.7258 .5288	.7721 .5976
.04	.5093 .2624	.5857 .3461	.6574 .4349	.7156 .5142	.7601 .5794	.5309 .2848	.6111 .3763	.6829 .4688	.7389 .5478	.7806 .6106
.05	.5443 .2991	.6131 .3786	.6764 .4599	.7283 .5323	.7687 .5923	.5653 .3223	.6370 .4084	.7002 .4924	.7502 .5644	.7882 .6225
			(N = 24)					(N = 26)		
.01	.3484 .1255	.4943 .2496	.6302 .4009	.7171 .5165	.7728 .5981	.3691 .1404	.5260 .2817	.6580 .4362	.7390 .5401	.7900 .6254
.02	.4470 .2034	.5613 .3190	.6624 .4417	.7336 .5402	.7820 .6129	.4672 .2219	.5870 .3482	.6862 .4735	.7534 .5693	.7985 .6388
.03	.5074 .2607	.6033 .3672	.6862 .4733	.7472 .5600	.7904 .6259	.5267 .2805	.6259 .3947	.7075 .5028	.7655 .5875	.8060 .6507
.04	.5508 .3063	.6341 .4048	.7052 .4994	.7588 .5774	.7979 .6377	.5694 .3270	.6548 .4313	.7247 .5271	.7759 .6034	.8128 .6616
.05	.5846 .3444	.6584 .4359	.7209 .5216	.7689 .5926	.8047 .6486	.6025 .3655	.6777 .4615	.7392 .5481	.7851 .6176	.8189 .6715
			(N = 28)					(N = 30)		
.01	.3893 .1559	.5548 .3124	.6821 .4681	.7578 .5760	.8051 .6493	.4090 .1717	.5810 .3419	.7031 .4968	.7740 .6006	.8182 .6705
.02	.4863 .2401	.6103 .3759	.7070 .5021	.7705 .5952	.8127 .6615	.5085 .2580	.6315 .4019	.7254 .5282	.7855 .6153	.8250 .6815
.03	.5488 .2999	.6465 .4207	.7263 .5294	.7814 .6119	.8195 .6725	.5618 .3186	.6651 .4449	.7439 .5536	.7953 .6337	.8312 .6917
.04	.5866 .3468	.6735 .4559	.7420 .5523	.7909 .6267	.8256 .6827	.6026 .3657	.6905 .4789	.7573 .5750	.8048 .6475	.8366 .7010
.05	.6189 .3854	.6950 .4850	.7553 .5720	.7992 .6330	.8312 .6917	.6341 .4064	.7168 .5071	.7696 .5937	.8117 .6599	.8419 .7395

TABLE OF [illegible] FOR [illegible] = 70

MEAN SQUARED DEVIATIONS OF G

LAMBDA/MU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.1571	.1799	.2141	.2672	.3440	.1816	.2169	.2736	.3538	.4602
	.0263	.0347	.0493	.0769	.1256	.0349	.0500	.0797	.1362	.2146
.02	.2269	.2562	.2956	.3479	.4117	.2562	.2399	.3572	.4232	.5054
	.0536	.0684	.0912	.1259	.1752	.0691	.0932	.1321	.1894	.2604
.03	.2759	.3076	.3477	.3969	.4526	.3109	.3540	.4084	.4715	.5356
	.0785	.0976	.1247	.1620	.2096	.0993	.1286	.1709	.2267	.2913
.04	.3143	.3469	.3863	.4324	.4827	.3515	.3944	.4458	.5026	.5592
	.1013	.1233	.1525	.1911	.2373	.1262	.1588	.2024	.2565	.3164
.05	.3460	.3787	.4172	.4605	.5066	.3847	.4268	.4753	.5273	.5784
	.1222	.1464	.1776	.2159	.2605	.1506	.1853	.2294	.2816	.3379
			(N = 12)					(N = 14)		
.01	.2052	.2556	.3388	.4489	.5490	.2281	.2960	.4044	.5229	.6139
	.0443	.0690	.1206	.2078	.3064	.0545	.0919	.1695	.2787	.3806
.02	.2673	.3425	.4170	.5016	.5790	.3147	.3836	.4724	.5619	.6353
	.0851	.1210	.1756	.2564	.3392	.1018	.1511	.2278	.3198	.4067
.03	.3426	.3973	.4644	.5359	.6014	.3719	.4377	.5145	.5895	.6525
	.1201	.1613	.2198	.2912	.3651	.1411	.1952	.2685	.3509	.4286
.04	.3848	.4377	.4989	.5619	.6196	.4151	.4770	.5456	.6111	.6669
	.1508	.1949	.2525	.3192	.3870	.1751	.2308	.3011	.3764	.4473
.05	.4189	.4697	.5262	.5830	.6349	.4497	.5081	.5703	.6290	.6794
	.1782	.2237	.2802	.3430	.4059	.2049	.2612	.3282	.3903	.4639
			(N = 16)					(N = 18)		
.01	.2506	.3375	.4651	.5834	.6626	.2727	.3790	.5182	.6277	.7004
	.0656	.1187	.2220	.3422	.4419	.0774	.1487	.2736	.3974	.4929
.02	.3406	.4229	.5218	.6112	.6790	.3652	.4599	.5650	.6513	.7136
	.1190	.1829	.2765	.3770	.4635	.1365	.2156	.3230	.4270	.5112
.03	.3991	.4749	.5586	.6336	.6927	.4245	.5092	.5968	.6698	.7248
	.1622	.2291	.3155	.4043	.4820	.1832	.2628	.3594	.4511	.5271
.04	.4427	.5127	.5861	.6517	.7045	.4682	.5450	.6212	.6852	.7347
	.1986	.2661	.3466	.4273	.4983	.2220	.3001	.3887	.4717	.5415
.05	.4775	.5424	.6063	.6669	.7149	.5329	.5732	.6411	.6984	.7436
	.2307	.2971	.3727	.4471	.5129	.2555	.3313	.4134	.4898	.5544

TABLE OF RHO STAR AND OMEGA STAR SQUARED FOR N = 70

MEAN SQUARED DEVIATIONS OF G

LAMBDA/MU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.2945 .0900	.4194 .1811	.5633 .3217	.6649 .4869	.7306 .5357	.3160 .1834	.4576 .2145	.6014 .3655	.6955 .4365	.7552 .5716
.02	.3887 .1543	.4944 .2484	.6022 .3660	.6843 .4707	.7415 .5515	.4111 .1723	.5261 .2806	.6342 .4052	.7117 .5085	.7645 .5859
.03	.4482 .2039	.5405 .2954	.6298 .3994	.7000 .4921	.7511 .5657	.4704 .2283	.5689 .3267	.6585 .4361	.7253 .5279	.7728 .5985
.04	.4917 .2446	.5741 .3325	.6515 .4269	.7133 .5107	.7596 .5784	.5136 .2665	.6004 .3632	.6778 .4616	.7369 .5446	.7802 .6099
.05	.5261 .2794	.6007 .3634	.6693 .4502	.7248 .5270	.7672 .5899	.5474 .3021	.6253 .3934	.6939 .4834	.7471 .5597	.7869 .6203
			(N = 24)					(N = 26)		
.01	.3372 .1174	.4931 .2479	.6338 .4049	.7209 .5217	.7757 .6030	.3581 .1320	.5256 .2808	.6615 .4404	.7425 .5530	.7930 .6299
.02	.4325 .1904	.5551 .3117	.6619 .4407	.7349 .5418	.7838 .6155	.4529 .2084	.5815 .3414	.6860 .4729	.7547 .5711	.8002 .6413
.03	.4913 .2443	.5947 .3566	.6833 .4691	.7468 .5592	.7911 .6269	.5109 .2639	.6181 .3847	.7050 .4990	.7653 .5870	.8067 .6517
.04	.5338 .2875	.6241 .3920	.7007 .4929	.7571 .5746	.7976 .6372	.5527 .3081	.6455 .4190	.7207 .5211	.7745 .6011	.8126 .6612
.05	.5671 .3240	.6475 .4215	.7153 .5134	.7662 .5884	.8037 .6468	.5853 .3449	.6675 .4476	.7340 .5403	.7828 .6139	.8180 .6699
			(N = 28)					(N = 30)		
.01	.3786 .1472	.5550 .3122	.6854 .4722	.7609 .5805	.8079 .6536	.3986 .1629	.5816 .3421	.7062 .5008	.7769 .6049	.8207 .6743
.02	.4723 .2264	.6055 .3697	.7070 .5018	.7718 .5970	.8143 .6680	.4908 .2442	.6273 .3963	.7254 .5280	.7467 .6200	.8265 .6839
.03	.5293 .2830	.6393 .4112	.7241 .5260	.7813 .6116	.8201 .6734	.5466 .3016	.6585 .4359	.7410 .5507	.7952 .6333	.8318 .6926
.04	.5703 .3277	.6649 .4443	.7384 .5467	.7896 .6246	.8254 .6821	.5867 .3466	.6825 .4678	.7540 .5699	.8029 .6855	.8367 .7003
.05	.6021 .3647	.6855 .4718	.7506 .5643	.7971 .6364	.8304 .6903	.6176 .3838	.7018 .4942	.7653 .5869	.8096 .6567	.8412 .7082

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR M = 80

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 8)					(N = 10)		
.01	.1506 .0241	.1737 .0323	.2092 .0471	.2659 .0759	.3487 .1284	.1743 .0321	.2102 .0469	.2700 .0774	.3624 .1375	.4669 .2239
.02	.2173 .0491	.2469 .0635	.2878 .0863	.3430 .1222	.4108 .1740	.2476 .0634	.2902 .0872	.3503 .1268	.4267 .1369	.5067 .2612
.03	.2643 .0720	.2965 .0906	.3381 .1177	.3899 .1561	.4491 .2062	.2982 .0912	.3425 .1203	.3996 .1635	.4665 .2217	.5343 .2894
.04	.3012 .0929	.3345 .1146	.3755 .1443	.4241 .1837	.4774 .2319	.3374 .1161	.3818 .1487	.4357 .1933	.4960 .2496	.5559 .3124
.05	.3318 .1123	.3654 .1362	.4054 .1675	.4512 .2071	.5001 .2537	.3696 .1390	.4133 .1736	.4644 .2189	.5196 .2733	.5738 .3323
			(N = 12)					(N = 14)		
.01	.1972 .0409	.2490 .0654	.3374 .1192	.4536 .2115	.5550 .3123	.2195 .0504	.2899 .0880	.4052 .1698	.5275 .2834	.6189 .3862
.02	.2759 .0784	.3328 .1142	.4115 .1737	.5009 .2553	.5810 .3412	.3027 .0941	.3744 .1439	.4684 .2237	.5623 .3199	.6374 .4091
.03	.3293 .1108	.3859 .1521	.4568 .2125	.5326 .2874	.6010 .3643	.3581 .1308	.4267 .1854	.5083 .2620	.5874 .3481	.6526 .4284
.04	.3702 .1395	.4252 .1838	.4900 .2435	.5570 .3134	.6175 .3841	.4000 .1625	.4650 .2192	.5379 .2924	.6074 .3717	.6656 .4452
.05	.4034 .1651	.4564 .2112	.5164 .2698	.5769 .3357	.6316 .4014	.4338 .1906	.4953 .2481	.5616 .3182	.6241 .3920	.6770 .4604
			(N = 16)					(N = 18)		
.01	.2415 .0608	.3323 .1149	.4674 .2237	.5859 .3470	.6668 .4471	.2633 .0720	.3750 .1454	.5210 .2759	.6317 .4020	.7040 .4976
.02	.3283 .1105	.4144 .1755	.5192 .2735	.6120 .3776	.6810 .4660	.3527 .1272	.4523 .2084	.5632 .3207	.6523 .4280	.7154 .5136
.03	.3850 .1508	.4647 .2192	.5535 .3097	.6322 .4023	.6930 .4822	.4102 .1710	.4997 .2529	.5927 .3542	.6689 .4496	.7253 .5277
.04	.4275 .1854	.5013 .2543	.5796 .3388	.6488 .4233	.7036 .4969	.4529 .2077	.5344 .2885	.6156 .3816	.6829 .4684	.7341 .5404
.05	.4615 .2154	.5303 .2839	.6007 .3634	.6629 .4415	.7130 .5101	.4869 .2395	.5617 .3181	.6344 .4048	.6951 .4850	.7421 .5521

TABLE OF RHO STAR AND OMEGA STAR SQUARE FOR N = 80

MEAN SQUARED DEVIATIONS OF G

LAMBDA/NU	.01	.02	.03	.04	.05	.01	.02	.03	.04	.05
			(N = 20)					(N = 22)		
.01	.2849	.4166	.5663	.6685	.7337	.3063	.4559	.6045	.6987	.7580
	.0842	.1784	.3246	.4493	.5399	.0970	.2125	.3687	.4902	.5759
.02	.3761	.4876	.6011	.6854	.7432	.3985	.5202	.6336	.7126	.7660
	.1444	.2415	.3644	.4719	.5538	.1618	.2741	.4041	.5099	.5880
.03	.4339	.5318	.6265	.6994	.7516	.4563	.5610	.6557	.7249	.7733
	.1910	.2859	.3951	.4911	.5662	.2109	.3176	.4322	.5271	.5991
.04	.4766	.5642	.6467	.7114	.7591	.4985	.5913	.6736	.7354	.7799
	.2296	.3210	.4205	.5078	.5774	.2510	.3521	.4557	.5423	.6092
.05	.5103	.5900	.6635	.7220	.7660	.5318	.6154	.6887	.7447	.7859
	.2628	.3505	.4422	.5228	.5880	.2851	.3809	.4761	.5559	.6186
			(N = 24)					(N = 26)		
.01	.3276	.4923	.6367	.7238	.7782	.3486	.5254	.6642	.7451	.7953
	.1107	.2468	.4083	.5256	.6067	.1250	.2801	.4437	.5567	.6335
.02	.4200	.5500	.6616	.7360	.7852	.4405	.5770	.6858	.7557	.8015
	.1794	.3058	.4400	.5432	.6175	.1971	.3360	.4724	.5724	.6433
.03	.4773	.5875	.6810	.7465	.7916	.4971	.6115	.7030	.7651	.8072
	.2305	.3479	.4658	.5587	.6276	.2498	.3764	.4960	.5866	.6524
.04	.5190	.6157	.6970	.7558	.7974	.5382	.6377	.7174	.7734	.8124
	.2719	.3814	.4876	.5725	.6367	.2921	.4089	.5163	.5992	.6608
.05	.5518	.6382	.7106	.7641	.8023	.5703	.6588	.7298	.7809	.8172
	.3067	.4094	.5066	.5850	.6453	.3274	.4359	.5340	.6108	.6685
			(N = 28)					(N = 30)		
.01	.3693	.5553	.6879	.7633	.8099	.3896	.5822	.7086	.7791	.8226
	.1400	.3122	.4753	.5839	.6567	.1555	.3424	.5340	.6081	.6774
.02	.4602	.6016	.7069	.7728	.8155	.4790	.6238	.7255	.7876	.8276
	.2148	.3647	.5015	.5983	.6658	.2324	.3917	.5280	.6213	.6856
.03	.5159	.6333	.7224	.7811	.8206	.5335	.6531	.7395	.7952	.8323
	.2688	.4034	.5235	.6111	.6741	.2872	.4287	.5483	.6332	.6933
.04	.5561	.6576	.7355	.7887	.8253	.5728	.6757	.7514	.8020	.8366
	.3115	.4344	.5424	.6230	.6818	.3304	.4585	.5659	.6440	.7005
.05	.5875	.6774	.7468	.7954	.8297	.6035	.6943	.7618	.8062	.8406
	.3473	.4607	.5590	.6336	.6890	.3662	.4837	.5814	.6540	.7072

A C T T E C H N I C A L B U L L E T I N N O . 16

Estimation of Proportions in Two-Way Tables

by

Ming-mei Wang

The American College Testing Program

and

Charles Lewis

The University of Illinois

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

September, 1973

1. Introduction

In two previous papers (Novick, Lewis, and Jackson, 1973; Lewis, Wang, and Novick, 1973), the problem of estimating proportions in m groups was studied with a Bayesian Model II approach, using the arc sine variance stabilizing transformation. It was shown that Bayesian Model II estimates were preferable to the conventional sample estimates especially when ϕ_Γ (the variance of the transformed variable γ_j) is small. This gain can be equated to substantial savings of sample size in data collection. An extension of this work is the problem of estimating proportions in two-way tables. For example, a set of t tests may be given to each of m persons. We are interested in estimating the level of functioning of each person on each test. By level of functioning on a test, we mean the percentage of correct responses that the person would make to a test composed of all of the items which might have been selected for the particular test. The model considered is the so-called Model II or random effects model because the persons and the tests are, respectively, considered to be random samples from larger populations of persons and tests. As in a two-way analysis of variance design, one can assume that the variations of performance are due to row effects (persons), column effects (tests), and interaction effects. Thus, each of these effects can be separately estimated and then combined to provide estimation of the proportions. This estimation procedure would find an application in the area of

---

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

individually prescribed instruction. After completing a unit of instruction, each student is given a posttest unit which consists of a set of tests on related skills. Estimates of the level of functioning of each student on each skill can be obtained to help make decisions on each individual's progress.

The Bayesian Model II approach was proposed by Lindley (1971). Theories and solutions for the general linear model have later been discussed in some detail by Lindley and Smith (1972). The present paper proposes to apply the Bayesian estimation procedures to two-way tables of proportions. Essentially, these procedures incorporate the collateral information provided by the other persons as well as by the other tests into the estimation of a single proportion. Consequently, some advantages are expected over the conventional sample proportion estimates.

2. Basic Model

The observed number of successes x_{ij} for individual i on test j is mapped into g_{ij} by the Freeman-Tukey (1950) transformation:

$$g_{ij} = \frac{1}{2}\left(\sin^{-1}\sqrt{\frac{x_{ij}}{n_j + 1}} + \sin^{-1}\sqrt{\frac{x_{ij} + 1}{n_j + 1}}\right) , \qquad (2.1)$$

where n_j is the number of items in test j . We will assume that the x_{ij} are binomially distributed with parameters n_j and π_{ij}, and that they are jointly independent given the π_{ij} . Under these assumptions, the g_{ij} are jointly independent and to a satisfactory approximation are normally distributed with mean $\gamma_{ij} = \sin^{-1}\sqrt{\pi_{ij}}$ and variance $v_{ij} = v_j = (4n_j + 2)^{-1}$, provided $n_j \geq 8$. The objective of this and

related transformations is variance stabilization. For further discussion on this point, as well as on the adequacy of the approximations, the reader is referred to Novick, Lewis, and Jackson, (1973), and Lewis, Wang, and Novick, (1973).

To proceed further, we must specify a distribution for Γ, the matrix of cell means γ_{ij}. If we treat the persons and tests as independent random samples from appropriate populations, then we may follow the standard development of random effects models given, for example, in Scheffé (1959, pp. 238-242). This development requires only that the persons and tests be sampled independently of each other and that the distribution of Γ be multivariate normal, given the necessary means and dispersion matrix. It is then possible to define θ, α_i, β_j, and δ_{ij} such that

$$\gamma_{ij} = \theta + \alpha_i + \beta_j + \delta_{ij} , \tag{2.2}$$

and such that $\{\alpha_i\}$, $\{\beta_j\}$, and $\{\delta_{ij}\}$ are independent normal with zero means and variances ϕ_α, ϕ_β, and ϕ_δ, respectively, conditional on these variances and independent of θ.

The definitions are given in terms of expectations of γ_{IJ} with respect to the population of persons and the population of tests. We indicate these expectations by $\mathcal{E}_I$ and $\mathcal{E}_J$, respectively:

$$\theta = \mathcal{E}_I \mathcal{E}_J(\gamma_{IJ}) , \tag{2.3}$$

$$\alpha_I = \mathcal{E}_J(\gamma_{IJ}) - \theta , \tag{2.4}$$

$$\beta_J = \mathcal{E}_I(\gamma_{IJ}) - \theta , \text{ and} \tag{2.5}$$

$$\delta_{IJ} = \gamma_{IJ} - \theta - \alpha_I - \beta_J . \tag{2.6}$$

In many cases where persons and tests have not been randomly sampled, it may, nonetheless, be possible to characterize our beliefs about the values of α_i as exchangeable for the group of persons being tested and for any other group selected from the population of interest. In addition, a similar statement may hold for β_j and the population of tests. Finally, our beliefs about the interaction terms (really residuals from a simple additive model) δ_{ij} may be exchangeable, at least in the sense that we have no good reason to expect any particular pattern of deviations from additivity in γ_{ij} .

Lindley and Smith (1972), among others, have applied the work of De Finetti (1937) and Hewitt and Savage (1955) on exchangeability to situations such as this. If we are willing to express our beliefs about γ_{ij} as described in the previous paragraph, we may conclude that $\{\alpha_i\}$ have the structure of identically and independently distributed random variables conditional on some parameter(s). Similar statements hold for $\{\beta_j\}$ and $\{\delta_{ij}\}$.

For mathematical convenience, we introduce the additional assumption that all the above-mentioned conditional distributions are normal. It immediately follows from definitions (2.4), (2.5), and (2.6) that the expectations of α_i, β_j, and δ_{ij} are zero. Hence, we may write

$$b(\underset{\sim}{\alpha}|\phi_\alpha) \propto \phi_\alpha^{-m/2} \exp(-\tfrac{1}{2}\,\phi_\alpha^{-1}\Sigma\alpha_i^2), \tag{2.7}$$

$$b(\underset{\sim}{\beta}|\phi_\beta) \propto \phi_\beta^{-t/2} \exp(-\tfrac{1}{2}\,\phi_\beta^{-1}\Sigma\beta_j^2), \tag{2.8}$$

and

$$b(\underset{\sim}{\Delta}|\phi_\delta) \propto \phi_\delta^{-mt/2} \exp(-\tfrac{1}{2}\,\phi_\delta^{-1}\Sigma\Sigma\delta_{ij}^2) \ . \tag{2.9}$$

A final assumption required at this stage is that $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and $\underset{\sim}{\Delta}$ are jointly independent, given ϕ_α, ϕ_β, and ϕ_δ . This assumption will be

reasonable as regards $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ whenever the choice of persons is unrelated to the choice of tests. The remainder of the assumption, namely the independence of $\underset{\sim}{\Delta}$ from $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ jointly, is less immediately intuitive but may be considered reasonable by noting that it is equivalent to the assertion that knowledge of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ tells us nothing about the distribution of $\underset{\sim}{\Delta}$, which may be a justifiable assertion on the basis of ignorance.

We have now reduced the problem of specifying a distribution for $\underset{\sim}{\Gamma}$ (either by standard methods or with suitable exchangeability and independence assumptions) to that of specifying θ, ϕ_α, ϕ_β, and ϕ_δ. In most cases, it will not be reasonable to assume that the values of these parameters are known. Consequently, we suggest the following distributional and independence assumptions: take ϕ_α, ϕ_β, and ϕ_δ to be independently distributed as inverse χ^2 variables and denote the degrees of freedom and sum of squares parameters for these distributions by $(\nu_\alpha, \lambda_\alpha)$, $(\nu_\beta, \lambda_\beta)$, and $(\nu_\delta, \lambda_\delta)$, respectively (see Novick and Jackson, 1974, Section 7.3). Finally, treat θ as locally uniform in the range of interest and jointly independent of ϕ_α, ϕ_β, ϕ_δ, $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and $\underset{\sim}{\Delta}$. We believe these distributions will satisfactorily characterize whatever vague knowledge we may have about the overall mean of γ_{ij} and its component variances. For reasons discussed by Novick (1969) and by Novick, Lewis, and Jackson (1973), a uniform distribution for θ will be acceptable; however, a proper prior will be required for the variances (ϕ_α, ϕ_β and ϕ_δ). Still outstanding is the issue of supplying values for the three pairs of inverse χ^2 parameters. We defer discussion on this point until Section 4.

With the above definitions and assumptions, we find the likelihood function and the joint posterior distribution of θ, $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, $\underset{\sim}{\Delta}$, ϕ_α, ϕ_β, and ϕ_δ given $\underset{\sim}{G} = (g_{ij})$ to be:

$$\ell(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Delta} \mid \underset{\sim}{G}) \propto \exp\{-\tfrac{1}{2} \sum_{ij}\sum (g_{ij} - \theta - \alpha_i - \beta_j - \delta_{ij})^2/v_j\}$$

$$i = 1, 2, \ldots, m; \; j = 1, 2, \ldots, t;$$

and

$$b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Delta}, \phi_\alpha, \phi_\beta, \phi_\delta | \underset{\sim}{G})$$

$$\propto \ell(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Delta} | \underset{\sim}{G}) \cdot b(\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Delta} | \phi_\alpha, \phi_\beta, \phi_\delta) \cdot b(\phi_\alpha, \phi_\beta, \phi_\delta)$$

$$\propto \exp\{-\tfrac{1}{2} \sum_{ij}\sum (g_{ij} - \theta - \alpha_i - \beta_j - \delta_{ij})^2/v_j\}$$

$$\phi_\alpha^{-\frac{1}{2}(m + \nu_\alpha + 2)} \exp\{-\tfrac{1}{2}(\lambda_\alpha + \sum_i \alpha_i^2)/\phi_\alpha\}$$

$$\phi_\beta^{-\frac{1}{2}(t + \nu_\beta + 2)} \exp\{-\tfrac{1}{2}(\lambda_\beta + \sum_j \beta_j^2)/\phi_\beta\}$$

$$\phi_\delta^{-\frac{1}{2}(mt + \nu_\delta + 2)} \exp\{-\tfrac{1}{2}(\lambda_\delta + \sum_{ij}\sum \delta_{ij}^2)/\phi_\delta\} , \quad (2.10)$$

respectively. We may use equation (2.2) to include $\underset{\sim}{\Gamma}$ explicitly in the joint posterior distribution (2.10). Specifically, we substitute $\gamma_{ij} - \theta - \alpha_i - \beta_j$ for δ_{ij} and leave the other parameters unchanged. Since the Jacobian of this transformation is unity, no further adjustments to (2.10) are necessary. Thus, we have

$$b(\underset{\sim}{\Gamma}, \theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \phi_\alpha, \phi_\beta, \phi_\delta | \underset{\sim}{G})$$

$$\propto \exp\{-\tfrac{1}{2} \sum_{ij}\sum (g_{ij} - \gamma_{ij})^2/v_j\} \cdot \phi_\alpha^{-\frac{1}{2}(m + \nu_\alpha + 2)} \exp\{-\tfrac{1}{2}(\lambda_\alpha + \sum_i \alpha_i^2)/\phi_\delta\}$$

$$\cdot\ \phi_\beta^{-\frac{1}{2}(t + \nu_\beta + 2)} \exp\{-\tfrac{1}{2}(\lambda_\beta + \sum_j \beta_j^2)/\phi_\beta\}$$

$$\cdot\ \phi_\delta^{-\frac{1}{2}(mt + \nu_\delta + 2)} \exp\{-\tfrac{1}{2}(\lambda_\delta + \sum_{ij}\sum \delta_{ij}^2)/\phi_\delta\} \quad . \qquad (2.11)$$

3. Posterior Joint Modal Estimates

3.1 Joint Modal Estimates for the Basic Model

Integrating $b(\underset{\sim}{\Gamma}, \theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \phi_\alpha, \phi_\beta, \phi_\delta | \underset{\sim}{G})$ in equation (2.11) with respect to the nuisance parameters ϕ_α, ϕ_β, and ϕ_δ, we obtain the posterior joint distribution of $\underset{\sim}{\Gamma}$, θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$:

$$b(\underset{\sim}{\Gamma}, \theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \underset{\sim}{G})$$

$$\propto \exp\{-\tfrac{1}{2} \sum_{ij}\sum (g_{ij} - \gamma_{ij})^2/v_j\} \cdot (\lambda_\alpha + \sum_i \alpha_i^2)^{-\frac{1}{2}(m + \nu_\alpha)}$$

$$\cdot (\lambda_\beta + \sum_i \beta_j^2)^{-\frac{1}{2}(t + \nu_\beta)} [\lambda_\delta + \sum_{ij}\sum (\gamma_{ij} - \theta - \alpha_i - \beta_j)^2]^{-\frac{1}{2}(mt + \nu_\delta)} \quad .$$

$$(3.1)$$

For the posterior joint distribution of $\underset{\sim}{\Gamma}$ alone, we need to integrate expression (3.1) with respect to θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ from equation (3.1). Explicit expressions for these integrations do not appear to us to be possible. Therefore, we obtain the joint mode of $\underset{\sim}{\Gamma}$, θ, $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ as estimates of the corresponding vector elements. Differentiating $f = \ell n\ b(\underset{\sim}{\Gamma}, \theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \underset{\sim}{G})$

with respect to θ, α_i, β_j, and γ_{ij} and setting the derivatives to zero, a system of equations is derived to be solved for the joint posterior mode of θ, $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and $\underset{\sim}{\Gamma}$:

$$\frac{\partial f}{\partial \gamma_{ij}} \quad : \quad v_j^{-1}(g_{ij} - \tilde{\gamma}_{ij}) - \tilde{\phi}_\delta^{-1}(\tilde{\gamma}_{ij} - \tilde{\theta} - \tilde{\alpha}_i - \tilde{\beta}_j) = 0 \ , \quad (3.2a)$$

$$\frac{\partial f}{\partial \theta} \quad : \quad \tilde{\phi}_\delta^{-1} \sum_{ij}\sum(\tilde{\gamma}_{ij} - \tilde{\theta} - \tilde{\alpha}_i - \tilde{\beta}_j) = 0 \ , \quad (3.2b)$$

$$\frac{\partial f}{\partial \alpha_i} \quad : \quad (\tilde{\phi}_\alpha^{-1} + t\tilde{\phi}_\delta^{-1})\tilde{\alpha}_i - t\tilde{\phi}_\delta^{-1}(\tilde{\gamma}_{i\cdot} - \tilde{\theta} - \tilde{\beta}_\cdot) = 0 \ , \quad (3.2c)$$

and

$$\frac{\partial f}{\partial \beta_j} \quad : \quad (\tilde{\phi}_\beta^{-1} + m\tilde{\phi}_\delta^{-1})\tilde{\beta}_j - m\tilde{\phi}_\delta^{-1}(\tilde{\gamma}_{\cdot j} - \tilde{\theta} - \tilde{\alpha}_\cdot) = 0 \ , \quad (3.2d)$$

where

$$\tilde{\phi}_\alpha = (\lambda_\alpha + \sum_i \tilde{\alpha}_i^2)/(m + \nu_\alpha) \ , \quad (3.3a)$$

$$\tilde{\phi}_\beta = (\lambda_\beta + \sum_j \tilde{\beta}_j^2)/(t + \nu_\beta) \ , \quad (3.3b)$$

and

$$\tilde{\phi}_\delta = [\lambda_\delta + \sum_{ij}\sum(\tilde{\gamma}_{ij} - \tilde{\theta} - \tilde{\alpha}_i - \tilde{\beta}_j)^2]/(mt + \nu_\delta) \ . \quad (3.3c)$$

Thus, we find the posterior joint modes:

$$\tilde{\gamma}_{ij} = \frac{\tilde{\phi}_\delta}{\tilde{\phi}_\delta + v_j} g_{ij} + \frac{v_j}{\tilde{\phi}_\delta + v_j} (\tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j) \ , \quad (3.4a)$$

$$\tilde{\alpha}_i = \omega_\alpha^*(\tilde{\gamma}_{i\cdot} - \tilde{\gamma}_{\cdot\cdot}) \ , \quad (3.4b)$$

$$\tilde{\beta}_j = \omega_\beta^*(\tilde{\gamma}_{\cdot j} - \tilde{\gamma}_{\cdot\cdot}) = \frac{\tilde{\phi}_\beta}{\tilde{\phi}_\beta + t^{-1}(\tilde{\phi}_\delta + v_j)} (g_{\cdot j} - \tilde{\gamma}_{\cdot\cdot}) \ , \quad (3.4c)$$

and

$$\tilde{\theta} = \tilde{\gamma}.. \ , \tag{3.4d}$$

where

$$\omega_\alpha^* = \frac{\tilde{\phi}_\alpha}{\tilde{\phi}_\alpha + t^{-1}\tilde{\phi}_\delta} \ , \quad \omega_\beta^* = \frac{\tilde{\phi}_\beta}{\tilde{\phi}_\beta + m^{-1}\tilde{\phi}_\delta} \ ,$$

and the dot notation indicates average over the appropriate index. Also, note that $\tilde{\alpha}. = \tilde{\beta}. = 0$.

The joint modal estimates of $\underset{\sim}{\Gamma}$, θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ then can be obtained by an iteration procedure. The usual least squares estimates $\hat{\theta} = g..$, $\hat{\alpha}_i = g_i. - g..$, $\hat{\beta}_j = g._j - g..$, and $\hat{\gamma}_{ij} = g_{ij}$ can be used as initial values of $\tilde{\theta}$, $\tilde{\alpha}_i$, $\tilde{\beta}_j$, and $\tilde{\gamma}_{ij}$. Given these initial values, $\tilde{\phi}_\alpha$, $\tilde{\phi}_\beta$, and $\tilde{\phi}_\delta$ are computed from (3.3) and used to obtain improved values of $\tilde{\underset{\sim}{\Gamma}}$, $\tilde{\underset{\sim}{\alpha}}$, $\tilde{\underset{\sim}{\beta}}$, and $\tilde{\theta}$ via (3.2). Substituting these new values in (3.3), the foregoing process is repeated to refine the estimates of $\tilde{\underset{\sim}{\Gamma}}$, $\tilde{\underset{\sim}{\alpha}}$, $\tilde{\underset{\sim}{\beta}}$, and $\tilde{\theta}$. This iterative procedure continues until some convergence criterion is reached. It should be noted that it may converge to some local mode if bimodality or multimodality exists.

Looking at the expressions in (3.4), we find that these Bayesian modal estimates $\tilde{\gamma}_{ij}$ are weighted averages of observed g_{ij} and the sum $\tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j$. In terms of (3.4b-d), it is seen that $\tilde{\theta}$, $\tilde{\alpha}_i$, and $\tilde{\beta}_j$ are functions of row averages $\tilde{\gamma}_i.$, column averages $\tilde{\gamma}._j$ and the overall average $\tilde{\gamma}..$. Under the basic model described in Section 2, the cell mean γ_{ij} for a specific cell (i, j) was assumed to be normally distributed with mean $\theta + \alpha_i + \beta_j$ and variance ϕ_δ, conditional on θ, α_i, β_j, and ϕ_δ . Thus, the weight $\frac{\tilde{\phi}_\delta}{\tilde{\phi}_\delta + v_j}$ assigned to the observed g_{ij} is a Bayesian

reliability estimate of g_{ij} given fixed i and j. . In obtaining estimates of γ_{ij}, observed g_{ij} are regressed towards the value $\tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j$. We may also remark that with the definition of γ_{ij} given by (2.2), $\gamma_{i\cdot} - \gamma_{\cdot\cdot}$ and $\gamma_{\cdot j} - \gamma_{\cdot\cdot}$ would be least squares estimates of α_i and β_j, respectively, provided γ_{ij} were observable. In this case, ω_α^* and ω_β^* can be regarded as their corresponding reliability estimates. Accordingly, the estimates $\tilde{\alpha}_i$ for α_i are seen to be $\tilde{\gamma}_{i\cdot} - \tilde{\gamma}_{\cdot\cdot}$ regressed towards their common value $\tilde{\alpha}_\cdot = 0$. This interpretation extends to the case of $\tilde{\beta}_j$.

Having obtained the posterior estimates of γ_{ij}, estimates of π_{ij} can be approximated by:

$$\tilde{\pi}_{ij} = (1 + \frac{1}{2n_j}) \sin^2 \tilde{\gamma}_{ij} - \frac{1}{4n_j} \quad , \tag{3.5}$$

following the previous study of the one-way case (Novick, Lewis, and Jackson, 1973, p. 24). In obtaining estimates $\tilde{\pi}_{ij}$ of proportions, the regression of sample proportions p_{ij} towards some common value corresponds to that of g_{ij} since $\tilde{\pi}_{ij}$ are monotonic increasing transformations of $\tilde{\gamma}_{ij}$.

3.2 A Special Case Where All Tests Are of Equal Length

In some applications, the set of related tests may have the same number of items, i.e., $n_j = n$. This means that the error variances within each cell are all equal ($v_j = v = \frac{1}{4n + 2}$). In this case, the solutions for (3.2) can be simplified. The joint modal estimates of $\underset{\sim}{\Gamma}$, θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ given $\underset{\sim}{G}$, can now be written as:

$$\begin{aligned}\tilde{\gamma}_{ij} &= \omega_\delta\, g_{ij} + (1 - \omega_\delta)(\tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j) \\ &= \omega_\delta\, g_{ij} + (1 - \omega_\delta)[g_{\cdot\cdot} + \omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + \omega_\beta(g_{\cdot j} - g_{\cdot\cdot})]\end{aligned} \tag{3.6a}$$

$$\tilde{\alpha}_i = \omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}) \tag{3.6b}$$

$$\tilde{\beta}_j = \omega_\beta(g_{\cdot j} - g_{\cdot\cdot}) \tag{3.6c}$$

and

$$\tilde{\theta} = g_{\cdot\cdot} = \sum_{ij}\sum g_{ij}/mt \tag{3.6d}$$

where

$$\omega_\delta = \frac{\tilde{\phi}_\delta}{\tilde{\phi}_\delta + v}\ ,\ \omega_\alpha = \frac{\tilde{\phi}_\alpha}{\tilde{\phi}_\alpha + t^{-1}(\tilde{\phi}_\delta + v)}\ ,\ \omega_\beta = \frac{\tilde{\phi}_\beta}{\tilde{\phi}_\beta + m^{-1}(\tilde{\phi}_\delta + v)} \tag{3.7}$$

and $\tilde{\phi}_\alpha$, $\tilde{\phi}_\beta$, $\tilde{\phi}_\delta$ are estimated from (3.3) .

Equation (3.6) indicates that the Bayesian Model II joint estimates of the transformed proportions γ_{ij} can be written explicitly as linear combinations of observed values g_{ij}, deviation row means $g_{i\cdot} - g_{\cdot\cdot}$, deviation column means $g_{\cdot j} - g_{\cdot\cdot}$, and the overall mean $g_{\cdot\cdot}$. The weights ω_α, ω_β, and ω_δ can be interpreted as reliability estimates of the components $g_{i\cdot} - g_{\cdot\cdot}$, $g_{\cdot j} - g_{\cdot\cdot}$, and g_{ij} . Consider θ, α_i, and β_j given, the basic assumptions in Section 2 imply

$$\mathrm{Var}(g_{ij} | \theta, \alpha_i, \beta_j) = \phi_\delta + v$$

and

$$\mathrm{Var}(\gamma_{ij} | \theta, \alpha_i, \beta_j) = \phi_\delta \ .$$

Hence, ω_δ is a reliability estimate of g_{ij}, conditional on θ, α_i, and β_j . Thus, it is seen that joint estimates of γ_{ij} are observed g_{ij} regressed towards $\tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j$.

The reliability interpretations of ω_α and ω_β may be less straightforward. However, borrowing from the results of classical random effect ANOVA, we obtain

$$\mathrm{Var}(g_{i\cdot} - g_{\cdot\cdot}) = \mathcal{E}\,\frac{\sum_i (g_{i\cdot} - g_{\cdot\cdot})^2}{m - 1} = \phi_\alpha + t^{-1}(\phi_\delta + v)\ ,$$

since $\sum_i (g_{i\cdot} - g_{\cdot\cdot})^2/(m - 1)$ is an unbiased estimate of $\mathrm{Var}(g_{i\cdot} - g_{\cdot\cdot})$ and $\mathcal{E}\ t\sum_i (g_{i\cdot} - g_{\cdot\cdot})^2/(m - 1) = t\phi_\alpha + \phi_\delta + v$ from the expected mean squares in random effect ANOVA. The sample statistic $g_{i\cdot} - g_{\cdot\cdot}$ is an estimate

of the row effect α_i, whose variance is assumed to be ϕ_α . Thus, ω_α can be regarded as an estimate of the reliability of $\hat{\alpha}_i = g_{i\cdot} - g_{\cdot\cdot}$. Since $\tilde{\alpha}_{\cdot} = 0$, we can write the joint modal estimates $\tilde{\alpha}_i$ of α_i in (3.6) as

$$\tilde{\alpha}_i = \omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + (1 - \omega_\alpha)\tilde{\alpha}_{\cdot} = \omega_\alpha\hat{\alpha}_i + (1 - \omega_\alpha)\tilde{\alpha}_{\cdot} \ .$$

It is clear then that the $\tilde{\alpha}_i$, being a weighted average of the least squares estimator $\hat{\alpha}_i$ and the common value $\tilde{\alpha}_{\cdot}$, are regressed towards $\tilde{\alpha}_{\cdot} = 0$. The same interpretation also extends to the case of the joint modal estimates $\tilde{\beta}_j$ of $\beta_j [\tilde{\beta}_j = \omega_\beta\hat{\beta}_j + (1 - \omega_\beta)\tilde{\beta}_{\cdot}]$.

Returning to (3.6), we can write

$$\tilde{\gamma}_{ij} = g_{ij} + (1 - \omega_\delta)[g_{\cdot\cdot} + \omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + \omega_\beta(g_{\cdot j} - g_{\cdot\cdot}) - g_{ij}] \ . \tag{3.8}$$

Hence, the regression of g_{ij} towards the estimate $\tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j =$ $g_{\cdot\cdot} + \omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + \omega_\beta(g_{\cdot j} - g_{\cdot\cdot})$ depends on the particular row i and column j . For instance, if the observed g_{ij} is greater than the value $\hat{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j$, γ_{ij} will be smaller than g_{ij} . The relative roles of a specific row i and column j in determining the direction of the regression of g_{ij} rest on the reliability estimates ω_α and ω_β . For example, if ω_α is much larger than ω_β, g_{ij} will be regressed mostly towards a combination of $g_{i\cdot}$ and $g_{\cdot\cdot}$.

In passing, it is also interesting to note that $\tilde{\gamma}_{i\cdot} = \sum_j \tilde{\gamma}_{ij}/t$ is a weighted average of $g_{i\cdot}$ and $g_{\cdot\cdot}$:

$$\tilde{\gamma}_{i\cdot} = [1 - (1 - \omega_\delta)(1 - \omega_\alpha)]g_{i\cdot} + (1 - \omega_\delta)(1 - \omega_\alpha)g_{\cdot\cdot}$$

$$= \omega_{\delta\alpha}g_{i\cdot} + (1 - \omega_{\delta\alpha})g_{\cdot\cdot} \ , \tag{3.9}$$

where

$$\omega_{\delta\alpha} = 1 - (1 - \omega_\delta)(1 - \omega_\alpha) = \frac{\tilde{\phi}_\alpha + t^{-1}\tilde{\phi}_\delta}{\tilde{\phi}_\alpha + t^{-1}(\tilde{\phi}_\delta + v)} \ .$$

$\omega_{\delta\alpha}$ is interpreted as an estimate of the reliability of $g_{i\cdot}$ conditional on θ and $\underset{\sim}{\beta}$ since $\mathrm{Var}(g_{i\cdot}|\theta, \underset{\sim}{\beta}) = \phi_\alpha + t^{-1}(\phi_\delta + v)$ and $\mathrm{Var}(\gamma_{i\cdot}|\theta, \underset{\sim}{\beta}) = \phi_\alpha + t^{-1}\phi_\delta$ under the assumptions discussed in Section 2. Therefore, estimates $\tilde{\gamma}_{i\cdot}$ of $\gamma_{i\cdot}$ are observed row averages $g_{i\cdot}$ regressed to the overall average $g_{\cdot\cdot}$. Similarly,

$$\tilde{\gamma}_{\cdot j} = \omega_{\delta\beta} g_{\cdot j} + (1 - \omega_{\delta\beta}) g_{\cdot\cdot} \tag{3.10}$$

where

$$\omega_{\delta\beta} = 1 - (1 - \omega_\delta)(1 - \omega_\beta) = \frac{\tilde{\phi}_\beta + m^{-1}\tilde{\phi}_\delta}{\tilde{\phi}_\beta + m^{-1}(\tilde{\phi}_\delta + v)}$$

is an estimate of the reliability of $g_{\cdot j}$ conditional on θ and $\underset{\sim}{\alpha}$.

3.3 A Generalized Case of the Present Model

Although we have discussed the problem of estimating proportions in two-way tables in the context of testing, this same model can be extended to a more general case where the indices n_{ij} of the binomial distributions for x_{ij} in cells (i, j) are all unequal. For example, one may be interested in simultaneously estimating proportions of female students in t different majors (Science, Art, etc.,) for each of m state universities. In this case, we may take samples of different sizes n_{ij} for each combination of majors and universities. Replacing all v_j by $v_{ij} = \frac{1}{4n_{ij} + 2}$ in (3.1), we obtain the posterior joint distributions of $\underset{\sim}{\Gamma}$, θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ for this general case. Thus, the joint modal estimates of $\underset{\sim}{\Gamma}$, θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ can be found by solving system (3.2) iteratively except substituting v_{ij} for v_j in (3.2a). The estimates $\tilde{\pi}_{ij}$ of π_{ij} are also obtained from (3.5) with n_{ij} substituting for n_j .

It may be noted that in this generalized application, our problem is similar to that treated by Lindley (1972) with two differences. Firstly, Lindley studies the general two-way ANOVA design so that there usually are replicated observations within each cell. In the present case, there is only one proportion observed for each cell. Secondly, we have a simpler case where all within (error) variances (v_{ij}) are known, while Lindley deals mostly with unknown within variances.

4. Further Discussion of the Prior Distributions

In Section 2, ϕ_α, ϕ_β, and ϕ_δ are apriori assumed to have independent inverse chi-square distributions with parameters $(\nu_\alpha, \lambda_\alpha)$, $(\nu_\beta, \lambda_\beta)$, and $(\nu_\delta, \lambda_\delta)$, respectively. In practice, the investigator must supply values for these three pairs of parameters to make the analysis feasible.

It has been argued by Novick, Lewis and Jackson, (1973), that in the absence of any specific information, a reasonable choice for the degrees of freedom parameter of the inverse chi-square distribution is 8 . If we accept this choice, the problem is reduced to that of specifying λ_α, λ_β, and λ_δ .

According to the assumptions made in Section 2, the prior marginal distribution of γ_{ij} conditional on θ, ϕ_α, ϕ_β, and ϕ_δ is normal with

$$\mathcal{E}(\gamma_{ij}|\theta, \phi_\alpha, \phi_\beta, \phi_\delta) = \theta$$

and

$$\mathrm{Var}(\gamma_{ij}|\theta, \phi_\alpha, \phi_\beta, \phi_\delta) = \phi_\alpha + \phi_\beta + \phi_\delta \quad . \tag{4.1}$$

Now if we ask the investigator about what the variance of γ_{ij} for a randomly chosen person-test combination (i, j), we obtain an estimate of the sum $\phi_\alpha + \phi_\beta + \phi_\delta$. This variance would usually be small, for example, between .04 and .02. We can further equate this estimate to the expected value of $\phi_\alpha + \phi_\beta + \phi_\delta$. Since independent inverse chi-square distributions are assumed for ϕ_α, ϕ_β, and ϕ_δ, we obtain

$$\mathcal{E}(\phi_\alpha + \phi_\beta + \phi_\delta) = \frac{\lambda_\alpha}{\nu_\alpha - 2} + \frac{\lambda_\beta}{\nu_\beta - 2} + \frac{\lambda_\delta}{\nu_\delta - 2} \quad . \tag{4.2}$$

Now, for illustrative purposes, assume the estimate of $\phi_\alpha + \phi_\beta + \phi_\delta$ is .02, i.e.,

$$\frac{\lambda_\alpha}{\nu_\alpha - 2} + \frac{\lambda_\beta}{\nu_\beta - 2} + \frac{\lambda_\delta}{\nu_\delta - 2} = .02 \ . \tag{4.3}$$

Combining (4.3) with the choice $\nu_\alpha = \nu_\beta = \nu_\delta = 8$, we should take

$$\nu_\alpha + \nu_\beta + \nu_\delta = .12 \ . \tag{4.4}$$

The investigator can now divide the total given in (4.4) among the three sum of squares parameters according to his prior beliefs as to the relative importance of person, test and person by test interaction effects on the transformed γ_{ij} . However, he should not set any of these parameters equal to zero for reasons discussed in Novick (1969). Thus, in the absence of specific information, he might choose

$$\lambda_\alpha = \lambda_\beta = \lambda_\delta = .04 \ . \tag{4.5}$$

In the next section, we shall examine, among other things, the effect of these choices on the estimates of γ_{ij} .

5. Numerical Examples

To illustrate the application of the present model, a set of data is constructed. There are 25 persons taking 5 related tests, each having 8 items. The observed proportions of correct answers are given in Table 1. These data were analyzed with different prior inverse chi-square distributions for ϕ_α, ϕ_β, and ϕ_δ. In Table 2, estimated proportions ($\tilde{\pi}_{ij}$) based on prior specifications $\nu_\alpha = \nu_\beta = \nu_\delta = 8$ and $\lambda_\alpha = \lambda_\beta = \lambda_\delta = .028$ are presented. It may be noted that since $\omega_\delta (= .0071)$ is negligible compared with ω_α and ω_β, the estimates $\tilde{\gamma}_{ij}$ are nearly completely regressed towards the combination $\omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + \omega_\beta(g_{\cdot j} - g_{\cdot\cdot}) + g_{\cdot\cdot} = \tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j$ (see equation 3.6a). Thus, the individually observed g_{ij} plays very little role in estimating γ_{ij} except through its contribution to $g_{i\cdot}$, $g_{\cdot j}$, and $g_{\cdot\cdot}$. Accordingly, the estimates $\tilde{\pi}_{ij}$ are largely dependent on the combined row, column, and overall averages of observed proportions.

In order to study the effects of prior parameters $(\nu_\alpha, \lambda_\alpha)$, $(\nu_\beta, \lambda_\beta)$, and $(\nu_\delta, \lambda_\delta)$ on the estimates $\tilde{\pi}_{ij}$, these data were also analyzed with $\nu_\alpha = \nu_\beta = \nu_\delta = 8$, $\lambda_\alpha = \lambda_\beta = \lambda_\delta = .06$, and $\nu_\alpha = \nu_\beta = \nu_\delta = 6$, $\lambda_\alpha = \lambda_\beta = \lambda_\delta = .10$. The results were presented in Tables 3 and 4, respectively. As can be seen from these tables, Bayesian estimates of ϕ_α, ϕ_β, and ϕ_δ are larger for bigger prior estimates ($\lambda_\alpha/\nu_\alpha$, λ_β/ν_β, and $\lambda_\delta/\nu_\delta$) of these variances. However, the increment of $\tilde{\phi}_\delta$ is smaller since its estimate is dominated by the sample information (with weight mt versus ν_δ). Consequently, as prior estimates of these variance components increase, ω_α and ω_β become comparatively larger while ω_δ does not change much. In general, there were not substantial differences among estimates $\tilde{\pi}_{ij}$ given in Tables 2, 3, and 4. They reveal the common

trend of regressing $\tilde{\gamma}_{ij}$ towards $\tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j$ and assigning very little weight to specific transformed cell proportions g_{ij}. As ω_α and ω_β increase, estimates $\tilde{\pi}_{ij}$ will increase for those (i, j) cells whose marginal averages $g_{i\cdot}$, $g_{\cdot j}$ are both bigger than the overall average $g_{\cdot\cdot}$. Conversely, if $g_{i\cdot}$ and $g_{\cdot j}$ are both smaller than $g_{\cdot\cdot}$, corresponding estimates $\tilde{\pi}_{ij}$ will be lower for higher ω_α and ω_β.

Finally, we may remark that for this data set, the classical estimate of ϕ_δ is negative. The sample statistic $\Sigma\Sigma(g_{ij} - g_{i\cdot} - g_{\cdot j} + g_{\cdot\cdot})^2$, whose expected value provides an estimate of $\phi_\delta + v$, is 2.1491. Therefore, the classical estimate of ϕ_δ based on expected mean squares is found to be $\hat{\phi}_\delta = [\Sigma\Sigma(g_{ij} - g_{i\cdot} - g_{\cdot j} + g_{\cdot\cdot})^2]/(m - 1)(t - 1) - v = -.0070$. For reference, we also calculated classical estimates of ϕ_α and ϕ_δ based on $\Sigma(g_{i\cdot} - g_{\cdot\cdot})^2 = .6878$ and $\Sigma(g_{\cdot j} - g_{\cdot\cdot})^2 = .00495$:

$$\hat{\phi}_\alpha = \frac{1}{t}[t\ \Sigma(g_{i\cdot} - g_{\cdot\cdot})^2/(m - 1) - \hat{\phi}_\delta - v] = .0242$$

and

$$\hat{\phi}_\beta = \frac{1}{m}[m\ \Sigma(g_{\cdot j} - g_{\cdot\cdot})^2/(t - 1) - \hat{\phi}_\delta - v] = .00034\ .$$

It is suspected that in the present context, the classical estimate of $\phi_\delta + v$ is based on only one observation per cell so that it is subject to large variations and thus highly unstable.

Table 1: Observed Proportions

$m = 25, t = 5, n = 8$

Subject/Test	1	2	3	4	5	Average	$g_{i\cdot}$
1	.875	.750	1.000	.750	.875	.850	1.149
2	.750	.625	.750	.500	.875	.700	.975
3	.875	1.000	1.000	.875	.875	.925	1.254
4	.750	.500	.625	.750	.750	.675	.947
5	.750	.875	.625	.750	1.000	.800	1.098
6	.875	.625	.750	.500	.625	.675	.951
7	1.000	.875	.875	1.000	1.000	.950	1.303
8	.875	.875	.750	.875	.625	.800	1.076
9	.750	.875	.750	1.000	.875	.850	1.149
10	.875	1.000	.625	1.000	.750	.850	1.175
11	.875	.875	.750	1.000	.875	.875	1.177
12	1.000	1.000	.875	.875	1.000	.950	1.303
13	1.000	1.000	1.000	1.000	1.000	1.000	1.401
14	1.000	1.000	.750	.875	1.000	.925	1.275
15	.750	1.000	.875	.625	.875	.825	1.126
16	.750	.875	.625	.750	.625	.725	.997
17	.500	.875	.625	.750	.625	.675	.951
18	.875	.375	.500	.625	.500	.575	.859
19	.500	.375	.375	.625	.750	.525	.809
20	.625	.500	.625	.500	.250	.500	.784
21	.750	1.000	.750	1.000	1.000	.900	1.248
22	.875	.875	1.000	.875	.750	.875	1.177
23	.750	.625	.625	.750	.625	.675	.946
24	.750	.875	.750	.500	.875	.750	1.026
25	1.000	1.000	.750	1.000	1.000	.950	1.324
Average	.815	.810	.745	.790	.800	.792	
$g_{\cdot j}$	1.115	1.129	1.039	1.101	1.112	$g_{\cdot\cdot} = 1.099$	

Table 2: Estimates of Proportions

$\nu_\alpha = \nu_\beta = \nu_\delta = 8$, $\lambda_\alpha = \lambda_\beta = \lambda_\delta = .028$

Subject/Test	1	2	3	4	5
1	.843	.850	.803	.835	.842
2	.762	.770	.714	.752	.761
3	.886	.894	.849	.879	.884
4	.748	.755	.699	.739	.746
5	.820	.829	.775	.812	.820
6	.751	.758	.702	.740	.747
7	.906	.911	.868	.899	.904
8	.811	.819	.765	.803	.807
9	.842	.851	.800	.837	.842
10	.854	.863	.811	.848	.852
11	.855	.862	.813	.849	.853
12	.906	.912	.868	.898	.904
13	.939	.945	.907	.934	.938
14	.896	.902	.856	.888	.894
15	.832	.842	.790	.824	.831
16	.773	.782	.725	.764	.770
17	.748	.760	.701	.741	.747
18	.704	.710	.651	.693	.699
19	.675	.684	.623	.666	.674
20	.662	.671	.611	.652	.658
21	.883	.892	.844	.878	.883
22	.855	.862	.815	.848	.853
23	.747	.755	.698	.739	.744
24	.787	.796	.740	.777	.786
25	.914	.920	.876	.907	.912

$\tilde{\phi}_\alpha = .0068$, $\tilde{\phi}_\beta = .0023$, $\tilde{\phi}_\delta = .0002$

$\omega_\alpha = .5357$, $\omega_\beta = .6620$, $\omega_\delta = .0071$

Table 3: Estimates of Proportions

$\nu_\alpha = \nu_\beta = \nu_\delta = 8, \quad \lambda_\alpha = \lambda_\beta = \lambda_\delta = .06$

Subject/Test	1	2	3	4	5
1	.849	.856	.801	.838	.847
2	.753	.762	.695	.739	.753
3	.897	.908	.855	.889	.895
4	.737	.745	.677	.727	.735
5	.821	.832	.766	.811	.824
6	.741	.748	.663	.725	.735
7	.920	.925	.875	.913	.919
8	.811	.821	.756	.802	.806
9	.847	.858	.796	.843	.847
10	.861	.873	.808	.855	.857
11	.862	.871	.811	.856	.860
12	.920	.928	.875	.910	.919
13	.956	.962	.920	.950	.955
14	.909	.917	.861	.899	.907
15	.835	.849	.785	.824	.835
16	.766	.778	.707	.756	.762
17	.736	.752	.679	.729	.735
18	.687	.691	.620	.672	.678
19	.650	.660	.586	.640	.651
20	.636	.646	.573	.622	.628
21	.893	.905	.847	.889	.895
22	.862	.871	.816	.853	.859
23	.736	.745	.676	.726	.732
24	.782	.794	.726	.769	.782
25	.929	.936	.884	.922	.927

$\phi_\alpha = .0100,\ \phi_\beta = .0049,\ \phi_\delta = .0005$

$\omega_\alpha = .6256,\ \omega_\beta = .8027,\ \omega_\delta = .0153$

Table 4: Estimates of Proportions

$\nu_\alpha = \nu_\beta = \nu_\delta = 6, \quad \lambda_\alpha = \lambda_\beta = \lambda_\delta = .10$

Subject/Test	1	2	3	4	5
1	.853	.860	.803	.840	.851
2	.746	.755	.682	.729	.747
3	.906	.918	.862	.897	.904
4	.728	.734	.660	.716	.725
5	.821	.834	.760	.811	.827
6	.734	.739	.666	.713	.725
7	.932	.935	.883	.924	.930
8	.812	.822	.749	.801	.803
9	.850	.863	.795	.848	.851
10	.866	.880	.807	.862	.861
11	.867	.877	.811	.863	.865
12	.932	.939	.883	.920	.930
13	.969	.975	.932	.963	.968
14	.920	.928	.866	.907	.918
15	.837	.855	.783	.824	.837
16	.760	.775	.694	.749	.755
17	.725	.745	.663	.719	.725
18	.672	.672	.595	.654	.660
19	.628	.638	.556	.618	.631
20	.613	.623	.544	.598	.601
21	.900	.915	.851	.899	.905
22	.867	.877	.819	.858	.863
23	.727	.736	.659	.715	.721
24	.778	.793	.717	.762	.779
25	.940	.948	.891	.933	.939

$\tilde{\phi}_\alpha = .0141$, $\tilde{\phi}_\beta = .0095$, $\tilde{\phi}_\delta = .0008$

$\omega_\alpha = .7001$, $\omega_\beta = .8866$, and $\omega_\delta = .0257$

REFERENCES

De Finetti, B. "Foresight: its logical laws, its subjective sources". In H. E. Kyburg, Jr., & G. E. Smokler (Eds.) Studies in Subjective Probability. New York: Wiley, Pp. 93-105, 1964. (The original French version appeared in Ann. de L'Inst. Henri Poincaré, 7, 1937.)

Freeman, M. F., & Tukey, J. W. "Transformations related to the angular and the square root". The Annals of Mathematical Statistics, 1950, 21, 607-611.

Hewitt, E., & Savage, L. J. Symmetric measures on cartesian products. Transactions of the American Mathematical Society, 1955, 80, 470-501.

Lewis, C., Wang, M., & Novick, M. R. Marginal distributions for the estimation of proportions in m groups. ACT Technical Bulletin No. 13. Iowa City, Iowa: The American College Testing Program, 1973.

Lindley, D. V. "The estimation of many parameters". In V. P. Godambe, & D. A. Sprott (Eds.) Foundations of Statistical Inference. Toronto: Holt, Rinehart, and Winston, 435-455, 1971.

Lindley, D. V. A Bayesian solution for two-way analysis of variance. ACT Technical Bulletin No. 8. Iowa City, Iowa: The American College Testing Program, 1972.

Lindley, D. V., & Smith, A. F. M. Bayesian estimates for the linear model. Journal of the Royal Statistical Society (Series B), 1972, 34, 1-41.

Novick, M. R. Multiparameter Bayesian indifference procedure. Journal of the Royal Statistical Society (Series B), 1969, 31, 29-64.

Novick, M. R., & Jackson, P. H. Statistical Methods for Educational and Psychological Research. New York: McGraw-Hill, 1974.

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-46.

Scheffe, H. The Analysis of Variance. New York: Wiley, 1959.

ACT TECHNICAL BULLETIN NO. 19

Marginal Distributions for the Estimation of Proportions in Two-Way Tables

by

Ming-mei Wang

The American College Testing Program

and

Charles Lewis

University of Illinois

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

December, 1973

Marginal Distributions for the Estimation of Proportions in Two-Way Tables

by

Ming-mei Wang

The American College Testing Program

and

Charles Lewis

University of Illinois

1. Introduction

The Bayesian Model II technique for simultaneous estimation of proportions in m groups has recently been extended to provide estimates of proportions in two-way tables by Wang and Lewis (1973). The random effects analysis of variance technique is applied to the transformed (observed) proportions

$$g_{ij} = \tfrac{1}{2}\left(\sin^{-1}\sqrt{\frac{x_{ij}}{n_j + 1}} + \sin^{-1}\sqrt{\frac{x_{ij} + 1}{n_j + 1}}\right),$$

where x_{ij} is the observed number of correct answers for person i on test j of n_j items (Freeman and Tukey, 1950). The g_{ij} are then assumed to be approximately normally distributed with mean γ_{ij} (= $\sin^{-1}\sqrt{\pi_{ij}}$, where the π_{ij} are the true proportions of successes for person i on test j) and known variance $v_j = (4n_j + 2)^{-1}$. The next step is to express γ_{ij} as a sum of the

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

overall effect θ, person effect α_i, test effect β_j, and person by test interaction effect δ_{ij} (these effects are defined in terms of expectations of γ_{ij} with respect to the appropriate populations of persons and tests, see Wang and Lewis, 1973, Section 2). Applying the exchangeability theorem (De Finetti, 1937), we obtain estimates of γ_{ij} by the Bayesian Model II procedure which incorporates not only the information provided by the performance of all persons on the test j, but also the information contained in the performance of the person i on all other tests into estimation of a single γ_{ij} . The resulting Bayesian estimates exhibit a regression of the least squares estimates $\hat{\alpha}_i = g_{i\cdot} - g_{\cdot\cdot}$, $\hat{\beta}_j = g_{\cdot j} - g_{\cdot\cdot}$, and $\hat{\delta}_{ij} = g_{ij} - g_{i\cdot} - g_{\cdot j} + g_{\cdot\cdot}$ towards their respective averages $\hat{\alpha}_{\cdot}$, $\hat{\beta}_{\cdot}$, and $\hat{\delta}_{\cdot\cdot}$, which, in the linear model, are each zero. Because we are making use of collateral information (provided by other tests as well as other persons) in estimating a specific proportion π_{ij}, it is expected, as in the one way m-group proportion case (Novick, Lewis, and Jackson, 1973), that some advantage will be gained over conventional methods.

The earlier paper (Wang and Lewis, 1973) provides us with the joint modal estimates of γ_{ij}, α_i, β_j, and θ . However, in applications to decision making in the context of individually prescribed instruction (IPI), the posterior marginal distributions of the γ_{ij} will be more useful. What is required in such applications is, for each test, the aposteriori probability that a person's test score is larger than some prespecified level, and this is required for each person and each test. As in the case of estimating proportions in m groups, algebraic expressions in closed form for these marginal distributions do not seem to exist. Hence, the marginal distributions of the γ_{ij} will be studied numerically in the present paper. In particular, we shall attempt to apply the numerical methods

developed by Lewis, Wang, and Novick (1973) in connection with estimating m group proportions to o.. present problem. Again, we will assume that all tests concerned are of equal length ($n_j = n$ for all j) to retain certain mathematical simplicity.

In the ANOVA terminology, the model adopted in the previous paper is the so-called non-additive model which includes interaction effects. Three variance components ϕ_α, ϕ_β, and ϕ_δ of person, test, and interaction effects, respectively, are postulated to account for the variability of γ_{ij}. In some cases, an additive model, one which assumes no interaction effects, may be a satisfactory alternative to the more general non-additive model. Specifically, in the context of IPI, the students are tested on related skills after studying the prescribed materials on a subject. Thus, the posttest unit consists of a set of tests which are very similar. If the tests are sufficiently similar (approaching τ-equivalence in the transformed units) so that the interaction variance component ϕ_δ is negligibly small compared with the other variance components, it will be adequate to choose an additive model in our analysis (see discussions in Lord and Novick, 1968, Section 7.6). In obtaining joint modal estimates for γ_{ij}, α_i, β_j, and θ, we have found that the contribution of individual g_{ij} to $\tilde{\gamma}_{ij}$ is negligible in nearly all examples we analyzed. This indicates that the estimates of γ_{ij} obtained from an additive model will be very close to those provided by the non-additive model. We may thus hope the additive model to be adequate for these data. The advantage of assuming an additive model is that the computational problem in obtaining marginal estimates for γ_{ij} will be much easier to handle while the amount of computational effort (especially computer time) required for this same purpose in the non-additive case is beyond practicality. We might note that while the arc-sine transformation is primarily designed

for variance stabilization, it also has a strong normalizing effect, which we have relied upon, and also some tendency to yield additivity. Our findings in these present applications seem to confirm remarks along these lines made to us by J. W. Tukey. For these reasons, the additive case of estimating proportions in two-way tables will be briefly discussed next, followed by the topic of marginal distributions.

2. An Additive Model for the Estimation of Gammas

In this section, we discuss a model which assumes no interaction effect for γ_{ij}. Thus, we may formulate

$$\gamma_{ij} = \theta + \alpha_i + \beta_j \tag{2.1}$$

where $\theta = \mathcal{E}_i\mathcal{E}_j\gamma_{ij}$, $\alpha_i = \mathcal{E}_j\gamma_{ij} - \theta$, and $\beta_j = \mathcal{E}_i\gamma_{ij} - \theta$ are defined in the appropriate person and test populations. It is assumed that the person effects are exchangeable apriori among all persons in the population. Similarly, the test effects are assumed to be exchangeable in he test population. We further assume that the distributions of α_i and β_j are normal and independent of each other. Finally, the prior distribution of θ is assumed to be locally uniform within the range of interest, and independent of the distributions of α_i and β_j (see also discussions in Lindley and Smith, 1972, and Smith, 1973). Thus, we obtain the likelihood and conditional prior density of θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ as:

$$\ell(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}|\underset{\sim}{G}) \propto \exp\{-\frac{1}{2v}\sum_{ij}\sum (g_{ij} - \theta - \alpha_i - \beta_j)^2\} \tag{2.2}$$

and

$$b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}|\phi_\alpha, \phi_\beta) \propto \phi_\alpha^{-\frac{m}{2}} \exp\{-\Sigma\alpha_i^2/2\phi_\alpha\}\, \phi_\beta^{-\frac{t}{2}} \exp\{-\Sigma\beta_j^2/2\phi_\beta\}, \tag{2.3}$$

where $\underset{\sim}{G} = (g_{ij})$ contains the transformed (observed) proportions g_{ij} for m persons on t tests and the vectors $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ contain the elements α_i and β_j, respectively. The corresponding matrix for the elements γ_{ij} will be denoted $\underset{\sim}{\Gamma}$. The zero means for the distributions of α_i and β_j are justified from the definitions of these effects. Note also that we discuss only the case of equal length tests so that $v = (4n + 2)^{-1}$ is used in Equation (2.2). To complete the model for our analysis, we assume, as usual, independent inverse chi-square prior distributions for the variances

ϕ_α and ϕ_β, with parameters $(\nu_\alpha, \lambda_\alpha)$ and $(\nu_\beta, \lambda_\beta)$, respectively (see Novick and Jackson, 1974, Section 7.6). Thus,

$$b(\phi_\alpha, \phi_\beta) \propto \phi_\alpha^{-\frac{1}{2}(\nu_\alpha + 2)} \exp\{-\lambda_\alpha/2\phi_\alpha\} \, \phi_\beta^{-\frac{1}{2}(\nu_\beta + 2)} \exp\{-\lambda_\beta/2\phi_\beta\} \, . \quad (2.4)$$

Combining Equations (2.2) through (2.4), and integrating with respect to ϕ_α and ϕ_β, we derive the posterior joint density of θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ to be:

$$b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \underset{\sim}{G}) \propto [\lambda_\alpha + \sum_i \alpha_i^2]^{-\frac{1}{2}(m + \nu_\alpha)} [\lambda_\beta + \sum_j \beta_j^2]^{-\frac{1}{2}(t + \nu_\beta)}$$

$$\exp\{-\frac{1}{2v} \sum_{ij}\sum (g_{ij} - \theta - \alpha_i - \beta_j)^2\} \, . \quad (2.5)$$

Therefore, the <u>joint</u> modal estimates of θ, $\underset{\sim}{\alpha}$, and β are found to be:

$$\tilde{\theta} = g_{\cdot\cdot} \quad (2.6a)$$

$$\tilde{\alpha}_i = \omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}), \quad \tilde{\alpha}_\cdot = 0 \quad (2.6b)$$

$$\tilde{\beta}_j = \omega_\beta(g_{\cdot j} - g_{\cdot\cdot}), \quad \tilde{\beta}_\cdot = 0 \quad (2.6c)$$

where

$$\omega_\alpha = \tilde{\phi}_\alpha/(\tilde{\phi}_\alpha + t^{-1}v), \ \omega_\beta = \tilde{\phi}_\beta/(\tilde{\phi}_\beta + m^{-1}v) \quad (2.7)$$

and

$$\tilde{\phi}_\alpha = (\lambda_\alpha + \Sigma\tilde{\alpha}_i^2)/(m + \nu_\alpha), \ \tilde{\phi}_\beta = (\lambda_\beta + \Sigma\tilde{\beta}_j^2)/(t + \nu_\beta) \, . \quad (2.8)$$

In terms of Equation (2.1), we obtain estimates for the γ_{ij}:

$$\tilde{\gamma}_{ij} = \tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j = g_{\cdot\cdot} + \omega_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + \omega_\beta(g_{\cdot j} - g_{\cdot\cdot}) \, . \quad (2.9)$$

Estimates of proportions based on $\tilde{\gamma}_{ij}$ are obtained from the sine-squared transformations as described in our earlier paper.

It should be pointed out that the $\tilde{\gamma}_{ij}$ are the estimates of γ_{ij} based on the joint modal estimates $\tilde{\theta}$, $\tilde{\underset{\sim}{\alpha}}$, and $\tilde{\underset{\sim}{\beta}}$ which maximize the joint posterior density of θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$. These estimates, $\tilde{\gamma}_{ij}$, are not joint modal estimates of the γ_{ij} from the full joint posterior distribution of the γ_{ij} . The joint distribution of the mt variables γ_{ij}, in this case, is degenerate with actual dimensionality $m + t - 1$ ($< mt$). To see this, we recall that the definition in Equation (2.1), in effect, shows that there can be only $m = t - 1$ linearly independent rows (or columns) in the coefficient matrix which generates γ_{ij} from θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$. That means, there exists at least a suitable subset Ω^* of $m + t - 1$ elements of the set $\Omega = \{\gamma_{ij}\}$ such that all the other $(mt - m - t + 1)$ elements in the complementary subset $\Omega - \Omega^*$ can be expressed as linear combinations of the elements in Ω^*. For example, given γ_{11}, γ_{12}, γ_{13}, and γ_{21} for the case $m = 2$, $t = 3$, we can write:

$$\gamma_{22} = \theta + \alpha_2 + \beta_2 = \gamma_{21} + \gamma_{12} - \gamma_{11} ,$$

and

$$\gamma_{23} = \theta + \alpha_2 + \beta_3 = \gamma_{21} + \gamma_{13} - \gamma_{11} .$$

Hence, the joint density can be defined in a space of dimensionality $m + t - 1$ at most (i.e., for at most $m + t - 1$ of the elements γ_{ij}). However, if we take any suitable subset Ω^* of $m + t - 1$ elements from the whole set Ω, the joint modal estimates of the γ_{ij} contained in Ω^* and the variables $\alpha.$ and $\beta.$ are identical to those $\tilde{\gamma}_{ij}$ obtained from the joint distribution of θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$. This can be shown by applying Lemma 3.2.3 in Anderson (1958, p. 47) and noting that the set of variables

$\{\Omega^*, \alpha., \beta.\}$ is a one-to-one transformation of the set $\{\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}\}$. Thus, the joint mode of $\{\Omega^*, \alpha., \beta.\}$ can be written as:

$$\tilde{\gamma}_{ij} = \tilde{\theta} + \tilde{\alpha}_i + \tilde{\beta}_j ,$$

for all γ_{ij} contained in Ω^*, and

$$\tilde{\alpha}. = \Sigma\tilde{\alpha}_i = 0, \ \tilde{\beta}. = \Sigma\tilde{\beta}_j = 0 .$$

In the above example, we can state that the joint modes of γ_{11}, γ_{12}, γ_{13}, γ_{21}, $\alpha.$, and $\beta.$ are $\tilde{\gamma}_{11}$, $\tilde{\gamma}_{12}$, $\tilde{\gamma}_{13}$, $\tilde{\gamma}_{21}$, $\tilde{\alpha}. = 0$, and $\tilde{\beta}. = 0$ as having obtained from the joint distribution of θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$.

To illustrate how this additive modal approximates the more general non-additive model used in our previous paper, we have re-analyzed the s. ne data with the present procedure. The data were explained in Table 1 of Wang and Lewis (1973). The same prior parameters ($\nu_\alpha = 8$, $\lambda_\alpha = .028$) and ($\nu_\beta = 8$, $\lambda_\beta = .028$) used in the non-additive case were adopted for the present analysis. Estimates of proportions obtained in the additive case are presented in Table 1. On comparing the results given in this table with those given in Table 2 of Wang and Lewis (1973), it is clearly seen that there are practically no differences between the two sets of estimates of proportions obtained from additive and non-additive models.

3. Conditional Posterior Distribution for Gammas

3.1 Additive Case

The additive model described in Section 2 can be summarized by the following statements:

(i) $\underset{\sim}{g}|\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}$ has a multivariate normal distribution with mean vector $\underset{\sim}{\gamma} = \underset{\sim}{A}\underset{\sim}{\xi}$ and dispersion matrix $v\underset{\sim}{I}$, where $\underset{\sim}{\xi}' = (\theta, \underset{\sim}{\alpha}', \underset{\sim}{\beta}')$ is a vector of order m + t + 1, and

$$A = \left(\begin{array}{c|cccc|c} & \underset{\sim}{1}_t & & & 0 & \underset{\sim}{I}_t \\ & & \underset{\sim}{1}_t & & & \underset{\sim}{I}_t \\ \underset{\sim}{1}_{mt} & & & \ddots & & \vdots \\ & 0 & & & \underset{\sim}{1}_t & \underset{\sim}{I}_t \end{array}\right)$$

is an mt by m + t + 1 matrix, where 0 indicate zeros in the rest of the matrix.

Here $\underset{\sim}{g}'$ and $\underset{\sim}{\gamma}'$ are row vectors whose elements are the rows of $\underset{\sim}{G}$ and $\underset{\sim}{\Gamma}$, respectively. Note that the notation $\underset{\sim}{1}_k$ is used to denote a k x 1 column vector, all of whose elements are 1 and $\underset{\sim}{I}_k$ an identity matrix of order k . The likelihood of $\underset{\sim}{\xi}$ can then be written as

$$\ell(\underset{\sim}{\xi}|\underset{\sim}{g}) \propto \exp\{-\frac{1}{2v}(\underset{\sim}{g} - \underset{\sim}{A}\underset{\sim}{\xi})'(\underset{\sim}{g} - \underset{\sim}{A}\underset{\sim}{\xi})\} . \qquad (3.1)$$

(ii) The conditional prior distribution of $\underset{\sim}{\xi}$ given ϕ_α and ϕ_β may be written in the following form:

$$b(\underset{\sim}{\xi}|\phi_\alpha, \phi_\beta) \propto \exp\{-\tfrac{1}{2}[\Sigma\alpha_i^2/\phi_\alpha + \Sigma\beta_j^2/\phi_\beta]\}$$

$$= \exp\{-\tfrac{1}{2}\underset{\sim}{\xi}'\underset{\sim}{D}^-\underset{\sim}{\xi}\}, \qquad (3.2)$$

where $\underset{\sim}{D}^-$ is the diagonal matrix of order (m + t + 1) defined as

$$\underset{\sim}{D}^- = \begin{pmatrix} 0 & & 0 \\ & \phi_\alpha^{-1}\underset{\sim}{I}_m & \\ 0 & & \phi_\beta^{-1}\underset{\sim}{I}_t \end{pmatrix} .$$

Combining Equations (3.1) and (3.2), the posterior conditional distribution of $\underset{\sim}{\xi}$ given ϕ_α and ϕ_β is found to be:

$$
\begin{aligned}
& b(\underset{\sim}{\xi} \mid \phi_\alpha, \phi_\beta, \underset{\sim}{g}) \\
& \quad \propto \ell(\underset{\sim}{g} \mid \underset{\sim}{\xi}) \cdot b(\underset{\sim}{\xi} \mid \phi_\alpha, \phi_\beta) \\
& \quad \propto \exp\{-\frac{1}{2v}(\underset{\sim}{g} - \underset{\sim}{A}\underset{\sim}{\xi})'(\underset{\sim}{g} - \underset{\sim}{A}\underset{\sim}{\xi}) - \tfrac{1}{2}\underset{\sim}{\xi}'\underset{\sim}{D}^{-}\underset{\sim}{\xi}\} \\
& \quad \propto \exp\{-\frac{1}{2v}(\underset{\sim}{\xi} - \underset{\sim}{B}\underset{\sim}{A}'\underset{\sim}{g})'\underset{\sim}{B}^{-1}(\underset{\sim}{\xi} - \underset{\sim}{B}\underset{\sim}{A}'\underset{\sim}{g})\}\ , \qquad (3.3)
\end{aligned}
$$

where $\underset{\sim}{B}^{-1} = \underset{\sim}{A}'\underset{\sim}{A} + v\underset{\sim}{D}^{-}$ is a nonsingular square symmetric matrix of order $(m + t + 1)$ and can be explicitly expressed as:

$$
\underset{\sim}{B}^{-1} = \begin{bmatrix} mt & t\,\underset{\sim}{1}'_m & m\,\underset{\sim}{1}'_t \\ t\,\underset{\sim}{1}_m & (t + v/\phi_\alpha)\underset{\sim}{I}_m & \underset{\sim}{1}_m\underset{\sim}{1}'_t \\ m\,\underset{\sim}{1}_t & \underset{\sim}{1}_t\,\underset{\sim}{1}'_m & (m + v/\phi_\beta)\underset{\sim}{I}_t \end{bmatrix} . \qquad (3.4)
$$

From Equation (3.3), we recognize that the conditional distribution of $\underset{\sim}{\xi}$, given ϕ_α, ϕ_β, and $\underset{\sim}{g}$ is a multivariate normal with mean vector $\mathcal{E}(\underset{\sim}{\xi} \mid \phi_\alpha, \phi_\beta, \underset{\sim}{g}) = \underset{\sim}{B}\underset{\sim}{A}'\underset{\sim}{g}$ and dispersion matrix $\mathrm{Var}(\underset{\sim}{\xi} \mid \phi_\alpha, \phi_\beta, \underset{\sim}{g}) = v\underset{\sim}{B}$. Thus, the mean vector and dispersion matrix can be obtained if the matrix $\underset{\sim}{B}^{-1}$ is inverted. Since the inverse of $\underset{\sim}{B}^{-1}$ is not easy to obtain directly by examining Equation (3.4), we now try to find the posterior conditional mean vector and dispersion matrix of $\underset{\sim}{\xi}$ by first considering the conditional distributions of $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and θ separately.

In terms of Equations (2.2) and (2.3), we have

$$b(\underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{G})$$

$$\propto \int \ell(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \underset{\sim}{G})\; b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta)\, d\theta$$

$$\propto \exp\{-\frac{1}{2v} \sum_{ij}\sum (g_{ij} - g_{..} - \alpha_i + \alpha_. - \beta_j + \beta_.)^2 - \frac{1}{2\phi_\alpha} \sum_i \alpha_i^2 - \frac{1}{2\phi_\beta} \sum_j \beta_j^2\}$$

$$\propto \exp\{-\tfrac{1}{2}[v^{-1} t \sum_i (g_{i.} - g_{..} - \alpha_i + \alpha_.)^2 + \phi_\alpha^{-1} \sum_i \alpha_i^2]\}$$

$$\cdot \exp\{-\tfrac{1}{2}[v^{-1} m \sum_j (g_{.j} - g_{..} - \beta_j + \beta_.)^2 + \phi_\beta^{-1} \sum_j \beta_j^2]\} \; . \quad (3.5)$$

Thus, $b(\alpha, \beta | \phi_\alpha, \phi_\beta, \underset{\sim}{G})$ can be factored into two parts and each involves only $\underset{\sim}{\alpha}$ or $\underset{\sim}{\beta}$. This implies that the conditional posterior distribution of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$, given ϕ_α, ϕ_β, and $\underset{\sim}{G}$, are independent of each other. We then can write

$$b(\underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) = b(\underset{\sim}{\alpha} | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) \cdot b(\underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) \; . \quad (3.6)$$

We may now proceed to find the posterior conditional mean vector $\underset{\sim}{u}_\alpha$ and dispersion matrix $\underset{\sim}{C}_\alpha$ of $\underset{\sim}{\alpha} | \phi_\alpha, \phi_\beta, \underset{\sim}{G}$ by observing:

$$b(\underset{\sim}{\alpha} | \phi_\alpha, \phi_\beta, \underset{\sim}{G})$$

$$\propto \exp\{-\tfrac{1}{2}[v^{-1} t \sum_i (g_{i.} - g_{..} - \alpha_i + \alpha_.)^2 + \phi_\alpha^{-1} \Sigma \alpha_i^2]\}$$

$$\propto \exp\{-\frac{t}{2vR_\alpha} [\Sigma \alpha_i^2 - 2R_\alpha \Sigma \alpha_i (g_{i.} - g_{..}) - mR_\alpha \alpha_.^2 + R_\alpha \Sigma (g_{i.} - g_{..})^2]\}$$

$$\propto \exp\{-\frac{t}{2vR_\alpha} [\Sigma (\alpha_i - R_\alpha g_{i.} + R_\alpha g_{..})^2 - mR_\alpha\, \alpha_.^2]\}$$

$$= \exp\{-\tfrac{1}{2}(\underset{\sim}{\alpha} - \underset{\sim}{u}_\alpha)' \underset{\sim}{C}_\alpha^{-1} (\underset{\sim}{\alpha} - \underset{\sim}{u}_\alpha)\} \; , \quad (3.7a)$$

where

$$\underset{\sim}{u}'_\alpha = [R_\alpha(g_{1\cdot} - g_{\cdot\cdot}),\ R_\alpha(g_{2\cdot} - g_{\cdot\cdot}),\ \ldots,\ R_\alpha(g_{m\cdot} - g_{\cdot\cdot})] \ , \qquad (3.7b)$$

$$\underset{\sim}{C}_\alpha^{-1} = \frac{t}{vR_\alpha}\ [\underset{\sim}{I}_m - \frac{R_\alpha}{m}\ \underset{\sim}{1}_m\underset{\sim}{1}'_m] \ ,$$

and

$$R_\alpha = \phi_\alpha/(\phi_\alpha + t^{-1}v) \ . \qquad (3.7c)$$

It follows that $\underset{\sim}{\alpha}|\phi_\alpha,\ \phi_\beta,\ \underset{\sim}{G}$ has a multivariate normal distribution with mean vector $\underset{\sim}{u}_\alpha$ and dispersion matrix

$$\underset{\sim}{C}_\alpha = \frac{vR_\alpha}{t}\left[\underset{\sim}{I}_m + \frac{R_\alpha\underset{\sim}{1}_m\underset{\sim}{1}'_m}{m(1 - R_\alpha)}\right] \ . \qquad (3.7d)$$

Similarly, the posterior conditional distribution of $\underset{\sim}{\beta}|\phi_\alpha,\ \phi_\beta,\ \underset{\sim}{G}$ is found to be a multivariate normal with mean vector

$$\underset{\sim}{u}'_\beta = [R_\beta(g_{\cdot 1} - g_{\cdot\cdot}),\ R_\beta(g_{\cdot 2} - g_{\cdot\cdot}),\ \ldots,\ R_\beta(g_{\cdot t} - g_{\cdot\cdot})] \ , \qquad (3.8a)$$

and dispersion matrix

$$\underset{\sim}{C}_\beta = \frac{vR_\beta}{m}\left[\underset{\sim}{I}_t + \frac{R_\beta\underset{\sim}{1}_t\underset{\sim}{1}'_t}{t(1 - R_\beta)}\right] , \qquad (3.8b)$$

where

$$R_\beta = \phi_\beta/(\phi_\beta + m^{-1}v) \ . \qquad (3.8c)$$

In order to obtain the conditional mean of $\theta|\phi_\alpha,\ \phi_\beta,\ \underset{\sim}{G}$, it is easier to first consider the conditional mean of $\theta|\phi_\alpha,\ \phi_\beta,\ \underset{\sim}{G},\ \underset{\sim}{\alpha}$. Since

$$b(\theta,\ \underset{\sim}{\alpha}|\phi_\alpha,\ \phi_\beta,\ \underset{\sim}{G})$$

$$\propto \int \ell(\theta,\ \underset{\sim}{\alpha},\ \underset{\sim}{\beta}|\underset{\sim}{G})\ b(\theta,\ \underset{\sim}{\alpha},\ \underset{\sim}{\beta}|\phi_\alpha,\ \phi_\beta)\ d\underset{\sim}{\beta}$$

$$\propto \exp\{-\frac{1}{2\phi_\alpha}\sum_i \alpha_i^2 - \frac{t}{2v}\sum_i (g_{i\cdot}-g_{\cdot\cdot}-\alpha_i+\alpha_\cdot)^2 - \frac{m}{2(m\phi_\beta + v)}\sum_j (g_{\cdot j}-\theta-\alpha_\cdot)^2\} \ ,$$

we find for given $\underset{\sim}{\alpha}$,

$$b(\theta|\phi_\alpha, \phi_\beta, \underset{\sim}{G}, \underset{\sim}{\alpha})$$

$$\propto \exp\{-\frac{mt}{2(m\phi_\beta + v)}[\theta - (g_{..} - \alpha_.)]^2\} \quad .$$

This implies

$$\mathcal{E}(\theta|\phi_\alpha, \phi_\beta, \underset{\sim}{G}, \underset{\sim}{\alpha}) = g_{..} - \alpha_. \quad .$$

Furthermore, referring to Equation (3.7b), we have

$$\begin{aligned} \mathcal{E}_{\underset{\sim}{\alpha}}(\alpha_.|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) &= \frac{1}{m}\sum_i \mathcal{E}(\alpha_i|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) \\ &= \frac{R_\alpha}{m}\sum_i (g_{i.} - g_{..}) \\ &= 0 \quad . \end{aligned}$$

Hence,

$$\mathcal{E}(\theta|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) = \mathcal{E}_{\underset{\sim}{\alpha}}\mathcal{E}(\theta|\phi_\alpha, \phi_\beta, \underset{\sim}{G}, \underset{\sim}{\alpha}) = g_{..} \quad . \tag{3.9}$$

The results of Equations (3.7b), (3.8a), and (3.9) lead us to conclude:

$$\begin{aligned} \mathcal{E}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) &= \mathcal{E}(\theta + \alpha_i + \beta_j|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) \\ &= g_{..} + R_\alpha(g_{i.} - g_{..}) + R_\beta(g_{.j} - g_{..}) \quad . \end{aligned} \tag{3.10}$$

It may be remarked that the conditional posterior mean of γ_{ij} given ϕ_α, ϕ_β, and $\underset{\sim}{G}$ takes a similar form to that of $\tilde{\gamma}_{ij}$ expressed in Equation (2.9) with R_α, R_β replacing ω_α, ω_β, respectively.

To obtain the conditional variance of γ_{ij} given ϕ_α, ϕ_β, and $\underset{\sim}{G}$, we have to find the dispersion matrix $v\underset{\sim}{B}$ of $\underset{\sim}{\xi}' = (\theta, \underset{\sim}{\alpha}', \underset{\sim}{\beta}')$. For this purpose we make use of the result $\mathcal{E}(\underset{\sim}{\xi}|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) = \underset{\sim}{B}\underset{\sim}{A}'g$. Since

$$(\underset{\sim}{A}'\underset{\sim}{g})' = (mtg.., tg_1., tg_2., \ldots, tg_m., mg._1, \ldots, mg._t) ,$$

and

$$[\mathcal{E}(\underset{\sim}{\xi}|\phi_\alpha, \phi_\beta, \underset{\sim}{G})]'$$

$$= [g.., R_\alpha(g_1.-g..), \ldots, R_\alpha(g_m.-g..), R_\beta(g._1-g..), \ldots, R_\beta(g._t-g..)]$$

as indicated by Equations (3.7b), (3.8a), and (3.9), we may write:

$$\underset{\sim}{B}\begin{bmatrix} mtg.. \\ tg_1. \\ \vdots \\ tg_m. \\ mg._1 \\ \vdots \\ mg._t \end{bmatrix} = \begin{bmatrix} g.. \\ R_\alpha(g_1. - g..) \\ \vdots \\ R_\alpha(g_m. - g..) \\ R_\beta(g._1 - g..) \\ \vdots \\ R_\beta(g._t - g..) \end{bmatrix} . \tag{3.11}$$

The elements of the matrix $\underset{\sim}{B}$ are known except those of its first row and first column. Explicitly, if we denote

$$Cov(\theta, \alpha_i|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) = vd_i \qquad i = 1, 2, \ldots, m$$

$$Cov(\theta, \beta_j|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) = ve_j \qquad j = 1, 2, \ldots, t$$

and

$$Var(\theta|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) = va ,$$

it is easy to verify that

$$\underset{\sim}{B} = v^{-1}\, Var(\underset{\sim}{\xi}|\phi_\alpha, \phi_\beta, \underset{\sim}{G})$$

$$= \begin{bmatrix} a & \underset{\sim}{d}' & \underset{\sim}{e}' \\ \underset{\sim}{d} & \underset{\sim}{C}_\alpha/v & \underset{\sim}{0} \\ \underset{\sim}{e} & \underset{\sim}{0}' & \underset{\sim}{C}_\beta/v \end{bmatrix} , \tag{3.12}$$

where $\underset{\sim}{d}' = (d_1, \ldots, d_m)$, $\underset{\sim}{e}' = (e_1, \ldots, e_t)$, $\underset{\sim}{C}_\alpha$ and $\underset{\sim}{C}_\beta$ are given in Equations (3.7d) and (3.8b), respectively, and $\underset{\sim}{0}$ is an m x t null matrix. Note also that $\text{Cov}(\alpha_i, \beta_j | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) = 0$ because the distributions of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ given ϕ_α, ϕ_β, and $\underset{\sim}{G}$ are independent [see Equation (3.6)]. Entering all the known elements of $\underset{\sim}{B}$ as shown in Equation (3.12) into Equation (3.11), the unknown elements a, d_i, and e_j can then be found by solving:

$$d_i(mtg_{\cdot\cdot}) + \left(\frac{R_\alpha}{t}\right) tg_{i\cdot} + \frac{R_\alpha^2}{mt(1 - R_\alpha)} \sum_i tg_{i\cdot} = R_\alpha(g_{i\cdot} - g_{\cdot\cdot}) \ , \quad (3.13a)$$

$$e_j(mtg_{\cdot\cdot}) + \left(\frac{R_\beta}{m}\right) mg_{\cdot j} + \frac{R_\beta^2}{mt(1 - R_\beta)} \sum_j mg_{\cdot j} = R_\beta(g_{\cdot j} - g_{\cdot\cdot}) \ , \quad (3.13b)$$

and

$$a(mtg_{\cdot\cdot}) + t \sum_i d_i \, g_{i\cdot} + m \sum_j e_j \, g_{\cdot j} = g_{\cdot\cdot} \ . \quad (3.13c)$$

It is easy to verify that the solutions to Equation (3.13) are:

$$d_i = -\frac{R_\alpha}{mt(1 - R_\alpha)} = v^{-1} \text{Cov}(\theta, \alpha_i | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) \ , \quad (3.14a)$$

$$e_j = -\frac{R_\beta}{mt(1 - R_\beta)} = v^{-1} \text{Cov}(\theta, \beta_j | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) \ , \quad (3.14b)$$

and

$$a = \frac{1}{mt}\left[1 + \frac{R_\alpha}{1 - R_\alpha} + \frac{R_\beta}{1 - R_\beta}\right] = v^{-1} \text{Var}(\theta | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) \ . \quad (3.14c)$$

The posterior conditional variances of γ_{ij} given ϕ_α, ϕ_β, and $\underset{\sim}{G}$ are obtained from Equations (3.7d), (3.8b), and (3.14a) through (3.14c), and the fact $\text{Cov}(\alpha_i, \beta_j | \phi_\alpha, \phi_\beta, \underset{\sim}{G}) = 0$:

$$\mathrm{Var}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \underset{\sim}{G})$$

$$= \mathrm{Var}(\theta + \alpha_i + \beta_j|\phi_\alpha, \phi_\beta, \underset{\sim}{G})$$

$$= \mathrm{Var}(\theta|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) + \mathrm{Var}(\alpha_i|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) + \mathrm{Var}(\beta_j|\phi_\alpha, \phi_\beta, \underset{\sim}{G})$$

$$+ 2\,\mathrm{Cov}(\theta, \alpha_i|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) + 2\,\mathrm{Cov}(\theta, \beta_j|\phi_\alpha, \phi_\beta, \underset{\sim}{G}) + 2\,\mathrm{Cov}(\alpha_i, \beta_j|\phi_\alpha, \phi_\beta, \underset{\sim}{G})$$

$$= v\left[\frac{R_\alpha}{t} + \frac{R_\beta}{m} + \frac{1 - R_\alpha - R_\beta}{mt}\right] . \qquad (3.15)$$

As for the distributions of $\gamma_{ij}|\phi_\alpha, \phi_\beta, \underset{\sim}{G}$, they are each known to be normal with mean and variance given in Equations (3.10) and (3.15), respectively. This follows from the definition of the γ_{ij} as linear combinations of θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ [see Equation (2.1)], and the result that $\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta}|\phi_\alpha, \phi_\beta, \underset{\sim}{G}$ has a $(m + t + 1)$-variate normal distribution [see Equation (3.3)]. In passing, we may note that similar results of the conditional means of θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$ given ϕ_α, ϕ_β, and $\underset{\sim}{G}$ have also been derived by Lindley and Smith (1972, Section 3.1), in connection with a general two-factor design without interaction.

3.2 Non-additive Case

In the non-additive case, the joint distribution of $\underset{\sim}{\gamma}$ is nondegenerate with dimensionality mt . This can be seen by examining the dispersion matrix $\underset{\sim}{C}^*$ of $\underset{\sim}{\gamma}$ given θ, ϕ_α, ϕ_β, and ϕ_δ under the assumptions made in the earlier paper by Wang and Lewis (1973):

$$\underset{\sim}{C}^* = \left[\begin{array}{c|c|c|c} (\phi_\beta+\phi_\delta)\underset{\sim}{I}_t+\phi_\alpha\underset{\sim}{1}_t\underset{\sim}{1}_t' & \phi_\beta\underset{\sim}{I}_t & \cdots & \phi_\beta\underset{\sim}{I}_t \\ \hline \phi_\beta\underset{\sim}{I}_t & (\phi_\beta+\phi_\delta)\underset{\sim}{I}_t+\phi_\alpha\underset{\sim}{1}_t\underset{\sim}{1}_t' & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline \phi_\beta\underset{\sim}{I}_t & \cdots & \ddots & (\phi_\beta+\phi_\delta)\underset{\sim}{I}_t+\phi_\alpha\underset{\sim}{1}_t\underset{\sim}{1}_t' \end{array}\right] \quad (3.16)$$

.

The matrix $\underset{\sim}{C}^*$ is a nonsingular symmetric matrix of order mt . In contrast, the dispersion matrix $\underset{\sim}{C}$ of $\underset{\sim}{\gamma}$ given θ, ϕ_α, and ϕ_β in the additive case as summarized from the assumptions made in Section 2,

$$\underset{\sim}{C} = \left[\begin{array}{c|c|c|c} \phi_\beta\underset{\sim}{I}_t+\phi_\alpha\underset{\sim}{1}_t\underset{\sim}{1}_t' & \phi_\beta\underset{\sim}{I}_t & \cdots & \phi_\beta\underset{\sim}{I}_t \\ \hline \phi_\beta\underset{\sim}{I}_t & \phi_\beta\underset{\sim}{I}_t+\phi_\alpha\underset{\sim}{1}_t\underset{\sim}{1}_t' & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline \phi_\beta\underset{\sim}{I}_t & \cdots & \ddots & \phi_\beta\underset{\sim}{I}_t+\phi_\alpha\underset{\sim}{1}_t\underset{\sim}{1}_t' \end{array}\right] \quad (3.17)$$

,

is singular, of rank $m + t - 1$ ($< mt$). Hence, as mentioned earlier, the joint distribution of $\underset{\sim}{\gamma}$ in this case is degenerate with actual dimensionality $m + t - 1$.

As a result of the above distinction, the posterior conditional distribution of γ_{ij} given ϕ_α, ϕ_β, ϕ_δ, and $\underset{\sim}{G}$ in the non-additive case can be approached in a somewhat different way. Under the assumptions made, the prior distribution of $\underset{\sim}{\gamma}$ can be described in two stages:

(i) Given θ, ϕ_α, ϕ_β, and ϕ_δ, $\underset{\sim}{\gamma}$ is, apriori, assumed to have an mt-variate normal distribution with mean vector $\theta\underset{\sim}{1}$ and dispersion matrix $\underset{\sim}{C}^*$ displayed in Equation (3.16). Note that for

convenience, we have used $\underset{\sim}{1}$ without subscript to denote an mt x 1 vector of ones (viz, $\underset{\sim}{1} = \underset{\sim}{1}_{mt}$).

(ii) Apriori, θ has a uniform distribution.

Combining these two statements with the normal likelihood of $\underset{\sim}{\gamma}$ given g, it can be shown, after a little algebra, that:

$$b(\underset{\sim}{\gamma}|\phi_\alpha, \phi_\beta, \phi_\delta, g)$$

$$\propto \int \exp\{-\frac{1}{2v}(\underset{\sim}{\gamma} - g)'(\underset{\sim}{\gamma} - g) - \tfrac{1}{2}(\underset{\sim}{\gamma} - \theta\underset{\sim}{1})'\underset{\sim}{C}^{*-1}(\underset{\sim}{\gamma} - \theta\underset{\sim}{1})\}\, d\theta$$

$$\propto \exp\{-\tfrac{1}{2}(\underset{\sim}{\gamma} - v^{-1}\underset{\sim}{B}^* g)'\underset{\sim}{B}^{*-1}(\underset{\sim}{\gamma} - v^{-1}\underset{\sim}{B}^* g)\}\ , \tag{3.18}$$

where

$$\underset{\sim}{B}^{*-1} = v^{-1}\underset{\sim}{I}_{mt} + \underset{\sim}{C}^{*-1} - \underset{\sim}{C}^{*-1}\underset{\sim}{1}(\underset{\sim}{1}'\underset{\sim}{C}^{*-1}\underset{\sim}{1})^{-1}\underset{\sim}{1}'\underset{\sim}{C}^{*-1}$$

is an mt x mt nonsingular symmetric matrix. Thus, we conclude that the posterior conditional joint distribution of $\underset{\sim}{\gamma}$ given ϕ_α, ϕ_β, ϕ_δ, and $\underset{\sim}{G}$ is a multivariate normal with mean vector $\underset{\sim}{a}^* = v^{-1}\underset{\sim}{B}^* g$ and dispersion matrix $\underset{\sim}{B}^*$.

Knowing the distributional form of $\underset{\sim}{\gamma}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{G}$, we now proceed to find $\underset{\sim}{a}^*$ and $\underset{\sim}{B}^*$ without actually carrying out the matrix inversion of $\underset{\sim}{B}^{*-1}$. Applying a lemma to be given later in Section 4, and referring to Equation (2.11) in Wang and Lewis (1973, p. 7), we can derive by integrating the expression w.r.t. θ, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$:

$$b(\underset{\sim}{\gamma}|\phi_\alpha, \phi_\beta, \phi_\delta, g)$$

$$\propto \phi_\delta^{-\frac{1}{2}(m-1)(t-1)}(\phi_\alpha + t^{-1}\phi_\delta)^{-\frac{m-1}{2}}(\phi_\beta + m^{-1}\phi_\delta)^{-\frac{t-1}{2}}$$

$$\cdot \exp\{-\frac{1}{2v}\sum_{ij}\sum(g_{ij} - \gamma_{ij})^2 - \frac{1}{2\phi_\delta}\sum_{ij}\sum(\gamma_{ij} - \gamma_{i\cdot} - \gamma_{\cdot j} + \gamma_{\cdot\cdot})^2\}$$

$$\cdot \exp\{-\frac{1}{2(\phi_\alpha + t^{-1}\phi_\delta)}\sum_i(\gamma_{i\cdot} - \gamma_{\cdot\cdot})^2 - \frac{1}{2(\phi_\beta + m^{-1}\phi_\delta)}\sum_j(\gamma_{\cdot j} - \gamma_{\cdot\cdot})^2\}\ . \tag{3.19}$$

Since we have shown that $b(\underset{\sim}{\gamma}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{g})$ is a multivariate normal density, the conditional mean of γ_{ij} given ϕ_α, ϕ_β, ϕ_δ, and $\underset{\sim}{g}$ can be found by solving the set of equations:

$$\partial \ln b(\underset{\sim}{\gamma}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{g})/\partial\gamma_{ij} = 0$$

for γ_{ij} . After some lengthy algebraic manipulations, this procedure leads to the result

$$\mathcal{E}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{g}) = R_\delta^* g_{ij} + (1 - R_\delta^*)[R_\alpha^* g_{i\cdot} + R_\beta^* g_{\cdot j} + (1 - R_\alpha^* - R_\beta^*)g_{\cdot\cdot}], \tag{3.20}$$

where

$$R_\delta^* = \frac{\phi_\delta}{\phi_\delta + v}, \quad R_\alpha^* = \frac{\phi_\alpha}{\phi_\alpha + t^{-1}(\phi_\delta + v)}, \text{ and } R_\beta^* = \frac{\phi_\beta}{\phi_\beta + m^{-1}(\phi_\delta + v)} . \tag{3.21}$$

To obtain the conditional variance of γ_{ij} given ϕ_α, ϕ_β, ϕ_δ, and $\underset{\sim}{g}$, we make use of the result $\mathcal{E}(\underset{\sim}{\gamma}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{g}) = v^{-1}\underset{\sim}{B}^*\underset{\sim}{g}$. This implies that the mt diagonal elements of the matrix $v^{-1}\underset{\sim}{B}^*$ are the coefficients of g_{ij} in expressing $\mathcal{E}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{g})$ in terms of the elements of $\underset{\sim}{g}$. Thus, we find from examining Equation (3.20)

$$\mathrm{Var}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{g}) = v[R_\delta^* + (1 - R_\delta^*)(\frac{R_\alpha^*}{t} + \frac{R_\beta^*}{m} + \frac{1 - R_\alpha^* - R_\beta^*}{mt})] , \tag{3.22}$$

because Equation (3.18) indicates that the diagonal elements of the matrix $\underset{\sim}{B}^*$ are the variances of $\gamma_{ij}|\phi_\alpha, \phi_\beta, \phi_\delta, \underset{\sim}{g}$. In summary then, we have shown that the posterior marginal distributions of γ_{ij}, conditional on ϕ_α, ϕ_β, and ϕ_δ, are normal with means and variances given by Equations (3.20) and (3.22), respectively.

4. Posterior Marginal Mean Estimates for Gammas

Having obtained the conditional means and variances of γ_{ij} given the variance components (ϕ_α, ϕ_β, and ϕ_δ), we now proceed to find marginal mean estimates of gammas by the same procedure used in Lewis, Wang, and Novick (1973). We shall restrict ourselves primarily to a consideration of the additive case as we have not yet been able to develop a practicable numerical algorithm for the non-additive case.

4.1 Additive Case

4.1.1 Posterior Distributions for the Variance Components

To obtain the posterior joint distribution of ϕ_α, ϕ_β, consider the function

$$b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{g}) = \ell(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \underset{\sim}{g}) \cdot b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta),$$

then,

$$b(\theta, \underset{\sim}{\alpha} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})$$

$$= \int b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})\, d\underset{\sim}{\beta}$$

$$\propto (\phi_\beta + m^{-1}v)^{-\frac{t}{2}} \phi_\alpha^{-\frac{m}{2}} \exp\{-\frac{1}{2\phi_\alpha} \sum_i \alpha_i^2 - \frac{1}{2v} \sum_{ij}\sum (g_{ij} - g_{\cdot j} - \alpha_i + \alpha_\cdot)^2\}$$

$$\cdot \exp\{-\tfrac{1}{2} \sum_j (g_{\cdot j} - \theta - \alpha_\cdot)^2 / (\phi_\beta + m^{-1}v)\} . \qquad (4.1)$$

Thus,

$$b(\underset{\sim}{\alpha} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})$$

$$= \int b(\theta, \underset{\sim}{\alpha} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})\, d\theta$$

$$\propto (\phi_\beta + m^{-1}v)^{-\frac{t-1}{2}} \phi_\alpha^{-\frac{m}{2}} \exp\{-\frac{1}{2\phi_\alpha} \sum_i \alpha_i^2 - \frac{1}{2v} \sum_{ij}\sum (g_{ij} - g_{\cdot j} - \alpha_i + \alpha_\cdot)^2\}$$

$$\cdot \exp\{-\tfrac{1}{2} \sum_j (g_{\cdot j} - g_{\cdot\cdot})^2 / (\phi_\beta + m^{-1}v)\} . \qquad (4.2)$$

Also,

$$b(\alpha, \beta | \phi_\alpha, \phi_\beta, g)$$

$$= \int b(\theta, \alpha, \beta | \phi_\alpha, \phi_\beta, g)\, d\theta$$

$$\propto \phi_\alpha^{-\frac{m}{2}} \exp\{-\frac{1}{2\phi_\alpha} \sum_i \alpha_i^2\} \cdot \phi_\beta^{-\frac{t}{2}} \exp\{-\frac{1}{2\phi_\beta} \sum_j \beta_j^2\}$$

$$\cdot \cdot \exp\{-\frac{1}{2v} \sum_{ij}\sum (g_{ij} - g_{..} - \alpha_i + \alpha_. - \beta_j + \beta_.)^2 \quad . \qquad (4.3)$$

Since

$$b(\alpha | \phi_\alpha, \phi_\beta, g) = \int b(\alpha, \beta | \phi_\alpha, \phi_\beta, g)\, d\beta \ ,$$

and

$$\sum_{ij}\sum (g_{ij} - g_{..} - \alpha_i + \alpha_. - \beta_j + \beta_.)^2 = \sum_{ij}\sum (g_{ij} - g_{.j} - \alpha_i + \alpha_.)^2 + m\sum_j (g_{.j} - g_{..} - \beta_j + \beta_.)^2 \ ,$$

we find from Equations (4.2) and (4.3):

$$1 \propto \int (\phi_\beta + m^{-1}v)^{\frac{t-1}{2}} \phi_\beta^{-\frac{t}{2}} \exp\{-\frac{1}{2\phi_\beta} \sum_j \beta_j^2 - \frac{m}{2v} \sum_j (g_{.j} - g_{..} - \beta_j + \beta_.)^2\}$$

$$\cdot \exp(\tfrac{1}{2} \sum_j (g_{.j} - g_{..})^2 / (\phi_\beta + m^{-1}v)\, d\beta \ .$$

Rearranging the above equation, and making the replacements $\phi_\beta = x$, $v = c$, $m = k$, $t = \ell$, and $\eta' = (\eta_1, \eta_2, \ldots, \eta_\ell) = \beta'$, we arrive at the following lemma:

$$\int x^{-\frac{\ell}{2}} \exp\{-\frac{1}{2x} \sum_j \eta_j^2 - \frac{k}{2c} \sum_j (g_{.j} - g_{..} - \eta_j + \eta_.)^2\}\, d\eta$$

$$\propto (x + k^{-1}c)^{-\frac{\ell-1}{2}} \exp\{-\frac{1}{2(x + k^{-1}c)} \sum_j (g_{.j} - g_{..})^2\} \ . \qquad (4.4)$$

Applying the above lemma to Equation (4.2) and noting that

$$\sum_{ij}\sum(g_{ij}-g_{\cdot j}-\alpha_i+\alpha_\cdot)^2 = \sum_{ij}\sum(g_{ij}-g_{i\cdot}-g_{\cdot j}+g_{\cdot\cdot})^2 + t\sum_i(g_{i\cdot}-g_{\cdot\cdot}-\alpha_i+\alpha_\cdot)^2 ,$$

it is easy to verify

$$\iiint b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})\, d\theta\, d\underset{\sim}{\alpha}\, d\underset{\sim}{\beta}$$

$$= \int b(\underset{\sim}{\alpha} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})\, d\underset{\sim}{\alpha}$$

$$\propto (\phi_\beta + m^{-1}v)^{-\frac{t-1}{2}} (\phi_\alpha + t^{-1}v)^{-\frac{m-1}{2}}$$

$$\cdot \exp\{-\frac{1}{2(\phi_\beta + m^{-1}v)}\sum_j(g_{\cdot j} - g_{\cdot\cdot})^2 - \frac{1}{2(\phi_\alpha + t^{-1}v)}\sum_i(g_{i\cdot} - g_{\cdot\cdot})^2\} . \tag{4.5}$$

It is further observed:

$$b(\phi_\alpha, \phi_\beta | \underset{\sim}{g}) = \iiint b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{g}) \cdot b(\phi_\alpha, \phi_\beta)\, d\theta\, d\underset{\sim}{\alpha}\, d\underset{\sim}{\beta}$$

$$= b(\phi_\alpha, \phi_\beta) \iiint b(\theta, \underset{\sim}{\alpha}, \underset{\sim}{\beta} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})\, d\theta\, d\underset{\sim}{\alpha}\, d\underset{\sim}{\beta} , \tag{4.6}$$

where $b(\phi_\alpha, \phi_\beta)$ is given in Equation (2.4). Consequently, upon substituting the expressions in Equations (2.4) and (4.5) into Equation (4.6), we have shown

$$b(\phi_\alpha, \phi_\beta | \underset{\sim}{g}) \propto b(\phi_\alpha | \underset{\sim}{g}) \cdot b(\phi_\beta | \underset{\sim}{g}) , \tag{4.7}$$

where

$$b(\phi_\alpha | \underset{\sim}{g}) \propto (\phi_\alpha + t^{-1}v)^{-\frac{m-1}{2}} \cdot \phi_\alpha^{-\frac{1}{2}(\nu_\alpha + 2)} \cdot \exp\{-\frac{S_R}{2(\phi_\alpha + t^{-1}v)} - \frac{\lambda_\alpha}{2\phi_\alpha}\} , \tag{4.8a}$$

and

$$b(\phi_\beta | \underset{\sim}{g}) \propto (\phi_\beta + m^{-1}v)^{-\frac{t-1}{2}} \phi_\beta^{-\frac{1}{2}(\nu_\beta + 2)} \cdot \exp\{-\frac{S_C}{2(\phi_\beta + m^{-1}v)} - \frac{\lambda_\beta}{2\phi_\beta}\} , \tag{4.8b}$$

with $S_R = \sum_i (g_{i\cdot} - g_{\cdot\cdot})^2$ and $S_C = \sum_j (g_{\cdot j} - g_{\cdot\cdot})^2$. The fact that $b(\phi_\alpha, \phi_\beta|\underset{\sim}{g})$ can be factored into the product of $b(\phi_\alpha|\underset{\sim}{g})$ and $b(\phi_\beta|\underset{\sim}{g})$ shows that the posterior distributions of ϕ_α and ϕ_β are independent and their density functions are given in Equations (4.8a) and (4.8b), respectively. It may also be pointed out that the mathematical forms of $b(\phi_\alpha|\underset{\sim}{g})$ and $b(\phi_\beta|\underset{\sim}{g})$ are similar to that of $b(\phi_\Gamma|\underset{\sim}{g})$ in the m-group proportion case [Lewis, Wang, and Novick, 1973, Equation (2.2), p. 6].

4.1.2 Posterior Marginal Means and Variances of Gammas

Having discussed the conditional posterior distributions of γ_{ij} given ϕ_α, ϕ_β, and $\underset{\sim}{g}$ in Section 3.1 and the posterior distributions of ϕ_α, ϕ_β in Section 4.1.1, the posterior marginal mean $\tilde{\gamma}_{ij}$ of γ_{ij} can be readily computed as:

$$\tilde{\gamma}_{ij} = \mathcal{E}(\gamma_{ij}|\underset{\sim}{g}) = \mathcal{E}_{\phi_\alpha} \mathcal{E}_{\phi_\beta} \mathcal{E}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \underset{\sim}{g})$$

$$= \mathcal{E}_{\phi_\alpha} \mathcal{E}_{\phi_\beta} [g_{\cdot\cdot} + R_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + R_\beta(g_{\cdot j} - g_{\cdot\cdot})]$$

$$= g_{\cdot\cdot} + \rho_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + \rho_\beta(g_{\cdot j} - g_{\cdot\cdot}) \ , \qquad (4.9)$$

where $\rho_\alpha = \mathcal{E}_{\phi_\alpha} R_\alpha$, $\rho_\beta = \mathcal{E}_{\phi_\beta} R_\beta$, and R_α, R_β are defined in Equations (3.7c) and (3.8c), respectively.

The marginal variance of γ_{ij} can also be obtained by using the relation:

$$\mathrm{Var}(\gamma_{ij}|\underset{\sim}{g}) = \mathcal{E}_{\phi_\alpha} \mathcal{E}_{\phi_\beta} [\mathrm{Var}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \underset{\sim}{g})] + \underset{\phi_\alpha,\, \phi_\beta}{\mathrm{Var}} [\mathcal{E}(\gamma_{ij}|\phi_\alpha, \phi_\beta, \underset{\sim}{g})] \ , \qquad (4.10)$$

where the $\underset{\phi_\alpha,\, \phi_\beta}{\mathrm{Var}}$ notation is used to denote

$$\underset{x,\, y}{\mathrm{Var}} [\mathcal{E}(w|x, y)] = \iint [\mathcal{E}(w|x, y) - \mathcal{E}_x \mathcal{E}_y \mathcal{E}(w|x, y)]^2 f(x, y)\, dx\, dy \ .$$

From Equation (3.15), we obtain:

$$\mathcal{E}_{\phi_\alpha} \mathcal{E}_{\phi_\beta} [\mathrm{Var}(\gamma_{ij} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})]$$

$$= \mathcal{E}_{\phi_\alpha} \mathcal{E}_{\phi_\beta} [v(\frac{R_\alpha}{t} + \frac{R_\beta}{m} + \frac{1 - R_\alpha - R_\beta}{mt})]$$

$$= v[\frac{1}{mt} + \frac{m-1}{mt} \rho_\alpha + \frac{t-1}{mt} \rho_\beta] . \tag{4.11}$$

The second term in Equation (4.10) is also easy to obtain:

$$\underset{\phi_\alpha, \phi_\beta}{\mathrm{Var}} [\mathcal{E} (\gamma_{ij} | \phi_\alpha, \phi_\beta, \underset{\sim}{g})]$$

$$= \underset{\phi_\alpha, \phi_\beta}{\mathrm{Var}} [g_{\cdot\cdot} + R_\alpha(g_{i\cdot} - g_{\cdot\cdot}) + R_\beta(g_{\cdot j} - g_{\cdot\cdot})]$$

$$= (g_{i\cdot} - g_{\cdot\cdot})^2 \underset{\phi_\alpha}{\mathrm{Var}}\, R_\alpha + (g_{\cdot j} - g_{\cdot\cdot})^2 \underset{\phi_\beta}{\mathrm{Var}}\, R_\beta$$

$$= (g_{i\cdot} - g_{\cdot\cdot})^2 [\mathcal{E}_{\phi_\alpha} R_\alpha^2 - \rho_\alpha^2] + (g_{\cdot j} - g_{\cdot\cdot})^2 [\mathcal{E}_{\phi_\beta} R_\beta^2 - \rho_\beta^2] . \tag{4.12}$$

Hence, we have reduced our problem to numerical computations of the values of ρ_α, ρ_β, $\mathcal{E}_{\phi_\alpha} R_\alpha^2$, and $\mathcal{E}_{\phi_\beta} R_\beta^2$. The integration problem here is closely related to that dealt with by Lewis, Wang, and Novick (1973). Thus, the same integration algorithm described there can be adopted for the present applications.

In general, we are interested in providing estimates for the proportions π_{ij} . This objective can be accomplished by applying the sine-squared transformations to $\tilde{\tilde{\gamma}}_{ij}$:

$$\tilde{\tilde{\pi}}_{ij} = (1 + \frac{1}{2n}) \sin^2 \tilde{\tilde{\gamma}}_{ij} - \frac{1}{4n} , \tag{4.13}$$

(see Novick, Lewis, and Jackson, 1973). A numerical example will be given in Section 5 to illustrate the estimation procedure outlined in this section.

4.1.3 Approximations to Marginal Probabilities

In theory, the posterior cumulative probability that π_{ij} is less than or equal to some value π_o can be computed by

$$\begin{aligned}
&\text{prob}(\pi_{ij} \le \pi_o|\underset{\sim}{g}) \\
&= \text{prob}(\gamma_{ij} \le \gamma_o|\underset{\sim}{g}) \\
&= \int_{-\infty}^{\gamma_o} b(\gamma_{ij}|\underset{\sim}{g})d\gamma_{ij} \\
&= \int_0^\infty \int_0^\infty \int_{-\infty}^{\gamma_o} b(\gamma_{ij}|\phi_\alpha, \phi_\beta, \underset{\sim}{g})b(\phi_\alpha, \phi_\beta|\underset{\sim}{g})d\gamma_{ij}\, d\phi_\alpha\, d\phi_\beta \\
&= \int_0^\infty \int_0^\infty \text{prob}(z \le z_o)b(\phi_\alpha|\underset{\sim}{g}) \cdot b(\phi_\beta|\underset{\sim}{g})d\phi_\alpha\, d\phi_\beta\ , \qquad (4.14)
\end{aligned}$$

where $\gamma_o = \sin^{-1}\sqrt{\pi_o}$, z is a standard normal variate and

$$z_o = \frac{\gamma_o - g_{\cdot\cdot} - R_\alpha(g_{i\cdot} - g_{\cdot\cdot}) - R_\beta(g_{\cdot j} - g_{\cdot\cdot})}{\left[v\left(\frac{R_\alpha}{t} + \frac{R_\beta}{m} + \frac{1 - R_\alpha - R_\beta}{mt}\right)\right]^{\frac{1}{2}}} \ . \qquad (4.15)$$

However, in practice, it is very time-consuming (beyond reasonable time limit with the algorithms we have tried) to evaluate this probability. We thus suggest a less ideal approach which is an extension of the result in Lewis, Wang, and Novick (1973). There it was found that the posterior distribution of γ_j, given $\underset{\sim}{g}$, can be satisfactorily approximated by a normal distribution with mean and variance equal to the posterior marginal mean and variance of γ_j, respectively. We venture to generalize this normal approximation to the present case. That is,

$$\text{prob}\{z \leq [\gamma_o - \mathcal{E}(\gamma_{ij}|\underset{\sim}{g})]/[\text{Var}(\gamma_{ij}|\underset{\sim}{g})]^{\frac{1}{2}}\}$$

will be evaluated to approximate $\text{prob}(\pi_{ij} \leq \pi_o|\underset{\sim}{g})$. While this approximation may not be as accurate as in the original application to the m-group proportion case, it should be sufficiently precise for deciding whether a student should be advanced to the next unit in IPI, provided the probability being estimated is not in the extreme tails.

4.2 Non-additive Case

Following the same procedure employed in Section 4.1.1, the posterior joint distribution of ϕ_α, ϕ_β, and ϕ_δ for the non-additive model can be derived using Equation (4.4) given in the lemma:

$$b(\phi_\alpha, \phi_\beta, \phi_\delta|\underset{\sim}{g})$$

$$\propto (\phi_\delta + v)^{-\frac{1}{2}(m-1)(t-1)}[\phi_\alpha + t^{-1}(\phi_\delta + v)]^{-\frac{m-1}{2}} \cdot [\phi_\beta + m^{-1}(\phi_\delta + v)]^{-\frac{t-1}{2}}$$

$$\cdot \exp\left\{-\frac{S_I}{2(\phi_\delta+v)} - \frac{S_R}{2[\phi_\alpha+t^{-1}(\phi_\delta+v)]} - \frac{S_C}{2[\phi_\beta+m^{-1}(\phi_\delta+v)]}\right\} \cdot b(\phi_\alpha, \phi_\beta, \phi_\delta), \tag{4.16}$$

where

$$S_I = \sum_{ij}\sum(g_{ij} - g_{i\cdot} - g_{\cdot j} + g_{\cdot\cdot})^2,$$

$$S_R = \sum_i(g_{i\cdot} - g_{\cdot\cdot})^2,$$

$$S_C = \sum_j(g_{\cdot j} - g_{\cdot\cdot})^2,$$

and $b(\phi_\alpha, \phi_\beta, \phi_\delta)$ is a product of three independent inverse chi-square densities with parameters $(\nu_\alpha, \lambda_\alpha)$, $(\nu_\beta, \lambda_\beta)$, and $(\nu_\delta, \lambda_\delta)$, respectively. It may be noted that a similar result has also been obtained by Box and Tiao (1973, p. 331) in their discussions of random effects ANOVA.

In contrast to the case of additive model, it is found from the Expression (4.16) that the posterior distributions of the three variance components are not independent of one another. Consequently, triple integrations are required to obtain posterior marginal means and variances for γ_{ij} . It appears, from our empirical experience, that the computer time needed for a triple integration of a function of the form in Equation (4.16) is at least the cube of what is needed for a simple integration of the function of the form in Equation (4.8a) or (4.8b). Unless some efficient approximations to these triple integrals can be found, this technique will not be practical for applications. Since we have not been able to devise an efficient algorithm which would complete this analysis with reasonable cost, we will not further discuss it.

5. A Numerical Example for the Additive Model

The data presented in Table 1 of an earlier paper (Wang and Lewis, 1973) are used for illustrative purposes here. Again, we choose $\nu_\alpha = \nu_\beta = 8$ and $\lambda_\alpha = \lambda_\beta = .028$ to characterize our prior distributions. There are 25(=m) persons, 5(=t) related tests and each test consists of 8(=n) items. As indicated in Sections 1 and 2, we feel that it is not far-fetched to assume an additive model for the analysis of these data.

The posterior marginal means and standard deviations of γ_{ij} given $\underset{\sim}{g}$ obtained from the procedures described in Section 4.1 are given in Table 2 (the figures enclosed in parentheses are standard deviations). It is found that both $\rho_\alpha = .7157$ and $\rho_\beta = .7140$ (weights used in computing marginal mean estimates $\tilde{\tilde{\gamma}}_{ij}$ of γ_{ij}) are larger than $\omega_\alpha = .5444$ and $\omega_\beta = .6637$ (the corresponding weights for obtaining the estimates $\tilde{\gamma}_{ij}$ of γ_{ij} based on posterior joint modes of $\underset{\sim}{\theta}$, $\underset{\sim}{\alpha}$, and $\underset{\sim}{\beta}$), respectively. From Equations (2.6b), (2.6c), (3.7b), and (3.8a), we find that marginal mean estimates of α_i and β_j are accordingly less regressed to their averages (zero) than the joint modal estimates. The smaller regressions of α_i and β_j in this case result in discrepancies between $\tilde{\gamma}_{ij}$ and $\tilde{\tilde{\gamma}}_{ij}$. The directions of these discrepancies depend on the signs of estimated person effect and test effect (which, in turn, are decided by the signs of $g_{i\cdot} - g_{\cdot\cdot}$ and $g_{\cdot j} - g_{\cdot\cdot}$). Specifically, if both $g_{i\cdot} - g_{\cdot\cdot}$ and $g_{\cdot j} - g_{\cdot\cdot}$ are positive (or negative), $\tilde{\tilde{\gamma}}_{ij}$ will be larger (or smaller) than $\tilde{\gamma}_{ij}$. On the other hand, if $g_{i\cdot} - g_{\cdot\cdot}$ and $g_{\cdot j} - g_{\cdot\cdot}$ are of opposite signs, their relative absolute values will decide the direction of the discrepancy and no general conclusions can be made.

The estimates $\tilde{\tilde{\pi}}_{ij}$ of proportions π_{ij} based on marginal estimates $\tilde{\tilde{\gamma}}_{ij}$ of γ_{ij} are presented in Table 3. It is seen that there are sizeable discrepancies between some $\tilde{\pi}_{ij}$ (based on $\tilde{\gamma}_{ij}$) and $\tilde{\tilde{\pi}}_{ij}$. For instance, the

estimates of proportion $\pi_{20,\ 3}$ for person 20 on test 3 are $\tilde{\pi}_{ij}$ = .608 and $\tilde{\tilde{\pi}}_{ij}$ = .548. To explain this difference, it is noted that $g_{i\cdot}$ = .784 and $g_{\cdot j}$ = 1.039 (i = 20, j = 3) are both smaller than the overall average $g_{\cdot\cdot}$ = 1.099 in this case. Therefore, the estimated person effect and test effect are negative. It follows that $\tilde{\tilde{\pi}}_{ij}$ is considerably smaller than $\tilde{\pi}_{ij}$ because $\rho_\alpha > \omega_\alpha$ and $\rho_\beta > \omega_\beta$.

In closing, an example of applying the proposed normal approximation to marginal probabilities is given below. From Table 2, we find $\mathcal{E}(\gamma_{20,\ 3}|\underset{\sim}{g})$ = .831 and $[\mathrm{Var}(\gamma_{20,\ 3}|\underset{\sim}{g})]^{\frac{1}{2}}$ = .0739 for person 20 on test 3. Suppose we are interested in a criterion mastery level π_o = .70. Following explanations in Section 4.1.3, we obtain

$$\begin{aligned} &\mathrm{prob}(\pi_{20,\ 3} > .70) \\ &\quad = 1 - \mathrm{prob}(\gamma_{20,\ 3} \leq .991) \\ &\quad \doteq 1 - \mathrm{prob}(z \leq 2.165) \\ &\quad = .0152 \end{aligned}$$

$$\left(\gamma_o = \sin^{-1}\sqrt{\pi_o} = .991 \text{ and } \frac{\gamma_o - \mathcal{E}(\gamma_{ij}|\underset{\sim}{g})}{[\mathrm{Var}(\gamma_{ij}|\underset{\sim}{g})]^{\frac{1}{2}}} = \frac{.991 - .831}{.0739} = 2.165\right).$$

Both the estimates $\tilde{\tilde{\pi}}_{ij}$ = .548 and $\tilde{\pi}_{ij}$ = .608 are less than .70 and the posterior probability that $\pi_{20,\ 3}$ is greater than .70 is very small. Thus, for most reasonable loss ratios, the action would be to retain this student in the old unit of instruction. For reference, approximate posterior probabilities of π_{ij} > .70 given $\underset{\sim}{g}$ are presented in Table 4. In this table, we find for person 2, $\mathrm{prob}(\pi_{21} > .70|\underset{\sim}{g}) \doteq .666$. Thus, we may decide to advance him on the basis of a loss ratio 2/1. Inspections of

Table 1 and 3 tell us both $\tilde{\pi}_{21} = .761$ and $\tilde{\tilde{\pi}}_{21} = .742$ are above the criterion .70. As another example, we find, for person 23 on test 4, both $\tilde{\pi}_{23,\ 4} = .737$ and $\tilde{\tilde{\pi}}_{23,\ 4} = .712$ are above .70. But we have $\text{prob}(\pi_{23,\ 4} > .70|g) \approx .498$ as given in Table 4. Therefore, we would advance him if the loss ratio is about 1 while retain him for any loss ratio greater than 1.

Table 1

Estimates of Proportions with an Additive Model

$\lambda_\alpha = \lambda_\beta = .028 \qquad \nu_\alpha = \nu_\beta = 8$

Test

Subject	1	2	3	4	5
1	.844	.851	.801	.836	.842
2	.761	.770	.713	.752	.759
3	.887	.894	.849	.880	.886
4	.747	.756	.698	.738	.745
5	.820	.828	.776	.812	.819
6	.749	.758	.700	.740	.747
7	.906	.913	.870	.900	.905
8	.810	.818	.765	.802	.808
9	.844	.851	.801	.836	.842
10	.855	.862	.813	.847	.853
11	.856	.863	.814	.848	.854
12	.906	.913	.870	.900	.905
13	.940	.946	.908	.935	.939
14	.896	.902	.858	.889	.894
15	.833	.841	.789	.825	.831
16	.772	.781	.725	.764	.770
17	.749	.758	.700	.740	.747
18	.701	.710	.650	.692	.699
19	.674	.683	.622	.664	.671
20	.660	.669	.608	.650	.658
21	.885	.892	.846	.878	.883
22	.856	.863	.814	.848	.854
23	.746	.755	.697	.737	.744
24	.786	.795	.740	.778	.784
25	.914	.920	.878	.908	.913

$\Sigma(g_{i\cdot} - g_{\cdot\cdot})^2 = .68782$, $\Sigma(g_{\cdot j} - g_{\cdot\cdot})^2 = .00495$

$\ddot{\phi}_\alpha = .00703$, $\ddot{\phi}_\beta = .00232$, $\omega_\alpha = .5444$, and $\omega_\beta = .6637$

Table 2

Posterior Marginal Means and Standard Deviations of Gammas

Test

Subject	1	2	3	4	5
1	1.146(.0704)	1.157(.0705)	1.092(.0706)	1.136(.0704)	1.144(.0704)
2	1.021(.0709)	1.032(.0709)	.967(.0711)	1.011(.0709)	1.019(.0709)
3	1.221(.0712)	1.231(.0712)	1.167(.0714)	1.211(.0712)	1.219(.0712)
4	1.002(.0712)	1.012(.0712)	.947(.0713)	.992(.0712)	.999(.0712)
5	1.110(.0704)	1.120(.0704)	1.055(.0705)	1.100(.0703)	1.107(.0704)
6	1.004(.0711)	1.015(.0712)	.950(.0713)	.994(.0711)	1.002(.0711)
7	1.256(.0718)	1.266(.0718)	1.202(.0720)	1.246(.0718)	1.254(.0718)
8	1.094(.0704)	1.104(.0704)	1.040(.0706)	1.084(.0704)	1.092(.0704)
9	1.146(.0704)	1.157(.0705)	1.092(.0706)	1.136(.0704)	1.144(.0704)
10	1.164(.0706)	1.175(.0706)	1.111(.0707)	1.154(.0705)	1.162(.0706)
11	1.166(.0706)	1.176(.0706)	1.112(.0707)	1.156(.0706)	1.164(.0706)
12	1.256(.0718)	1.266(.0718)	1.202(.0720)	1.246(.0718)	1.254(.0718)
13	1.326(.0735)	1.337(.0735)	1.272(.0736)	1.316(.0735)	1.324(.0735)
14	1.236(.0714)	1.247(.0715)	1.182(.0716)	1.226(.0714)	1.234(.0714)
15	1.129(.0704)	1.140(.0704)	1.075(.0706)	1.119(.0704)	1.127(.0704)
16	1.038(.0707)	1.048(.0708)	.983(.0709)	1.028(.0707)	1.035(.0707)
17	1.004(.0711)	1.015(.0712)	.950(.0713)	.994(.0711)	1.002(.0711)
18	.939(.0723)	.949(.0724)	.885(.0725)	.929(.0723)	.937(.0723)
19	.903(.0732)	.913(.0733)	.849(.0734)	.893(.0732)	.901(.0732)
20	.885(.0738)	.895(.0738)	.831(.0739)	.875(.0738)	.883(.0737)
21	1.217(.0711)	1.227(.0711)	1.162(.0713)	1.207(.0711)	1.214(.0711)
22	1.166(.0706)	1.176(.0706)	1.112(.0707)	1.156(.0706)	1.164(.0706)
23	1.001(.0712)	1.011(.0712)	.946(.0713)	.991(.0712)	.998(.0712)
24	1.058(.0705)	1.069(.0706)	1.004(.0707)	1.048(.0705)	1.056(.0705)
25	1.272(.0721)	1.282(.0723)	1.217(.0723)	1.262(.0721)	1.269(.0721)

Standard deviations are given in parentheses. Prior Specifications: $\lambda_\alpha = \lambda_\beta = .028$, $\nu_\alpha = \nu_\beta = 8$; Sample Statistics, $\Sigma(g_{i\cdot} - g_{\cdot\cdot})^2 = .68782$, $\Sigma(g_{\cdot j} - g_{\cdot\cdot})^2 = .00495$; $\rho_\alpha = \mathcal{E}_{\phi_\alpha} R_\alpha = .7157$, $\rho_\beta = \mathcal{E}_{\phi_\beta} R_\beta = .7140$, $\underset{\phi_\alpha}{\mathrm{Var}}\, R_\alpha = .00492$, and $\underset{\phi_\beta}{\mathrm{Var}}\, R_\beta = .00712$.

Table 3

Posterior Marginal Estimates of Proportions Based on $\mathcal{E}(\gamma_{ij}|\underset{\sim}{g})$

	Test				
Subject	1	2	3	4	5
1	.851	.859	.806	.843	.849
2	.742	.751	.689	.732	.739
3	.907	.913	.867	.900	.905
4	.723	.733	.669	.713	.721
5	.821	.829	.773	.812	.819
6	.725	.735	.672	.716	.723
7	.930	.936	.893	.923	.928
8	.808	.816	.759	.799	.806
9	.851	.859	.806	.843	.849
10	.865	.873	.821	.858	.864
11	.867	.874	.823	.859	.865
12	.930	.936	.893	.923	.928
13	.969	.974	.939	.964	.968
14	.917	.924	.879	.910	.915
15	.837	.846	.791	.829	.835
16	.757	.766	.705	.747	.755
17	.726	.735	.672	.716	.723
18	.661	.671	.605	.650	.658
19	.624	.634	.567	.613	.621
20	.605	.616	.548	.595	.603
21	.904	.911	.864	.897	.902
22	.867	.874	.823	.859	.865
23	.722	.732	.668	.712	.720
24	.776	.785	.724	.767	.774
25	.939	.945	.904	.933	.938

Prior Specifications: $\lambda_\alpha = \lambda_\beta = .028$, $\nu_\alpha = \nu_\beta = 2$

Table 4

Approximate Posterior Probabilities of $\pi_{ij} > .70$

Test

Subject	1	2	3	4	5
1	.986	.991	.924	.980	.985
2	.666	.716	.368	.613	.654
3	.999	1.000	.993	.999	.999
4	.559	.615	.270	.503	.546
5	.954	.966	.818	.938	.951
6	.574	.629	.282	.518	.561
7	1.000	1.000	.999	1.000	1.000
8	.928	.946	.755	.907	.924
9	.986	.991	.924	.980	.985
10	.993	.995	.954	.990	.992
11	.993	.996	.956	.990	.993
12	1.000	1.000	1.000	1.000	1.000
13	1.000	1.000	1.000	1.000	1.000
14	1.000	1.000	.996	1.000	1.000
15	.975	.982	.883	.966	.973
16	.744	.789	.456	.697	.734
17	.574	.629	.282	.518	.561
18	.235	.281	.071	.195	.225
19	.114	.144	.026	.090	.108
20	.075	.097	.015	.058	.071
21	.999	1.000	.992	.999	.999
22	.993	.996	.956	.990	.993
23	.554	.610	.266	.498	.541
24	.829	.864	.572	.791	.821
25	1.000	1.000	.999	1.000	1.000

References

Anderson, T. W. Introduction to Multivariate Statistical Analysis. New York: Wiley, 1958.

Box, G. E. P., & Tiao, G. C. Bayesian Inference in Statistical Analysis. Reading, Mass.: Addison-Wesley, 1973.

De Finetti, B. Foresight: Its logical laws, its subjective sources. In H. E. Kyburg, Jr., & G. E. Smokler (Eds.). Studies in Subjective Probability. New York: Wiley, 1964, 93-105. (The original French version appeared in Ann. de L'Inst. Henri Poincaré, 1937, 7)

Freeman, M. F., & Tukey, J. W. Transformations related to the angular and the square root. The Annals of Mathematical Statistics, 1950, 21, 607-611.

Lewis, C., Wang, M., & Novick, M. R. Marginal distributions for the estimation of proportions in m groups. ACT Technical Bulletin No. 13. Iowa City, Iowa: The American College Testing Program, 1973.

Lord, F. M., & Novick, M. R. Statistical Theories of Mental Test Scores. Reading, Mass.: Addison-Wesley, 1968.

Lindley, D. V., & Smith, A. F. M. Bayesian estimates for the linear model. Journal of the Royal Statistical Society (Series B), 1972, 34, 1-41.

Novick, M. R., & Jackson, P. H. Statistical Methods for Educational and Psychological Research. New York: McGraw Hill, 1974.

Novick, M. R., Lewis, C., & Jackson, P. H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-46.

Smith, A. F. M. A general Bayesian linear model. Journal of the Royal Statistical Society (Series B), 1973, 35, 67-75.

Wang, M., & Lewis, C. Estimation of proportions in two-way tables. ACT Technical Bulletin No. 16. Iowa City, Iowa: The American College Testing Program, 1973.

Appendix Number 4

to

Final Report

Project No. 2-0067

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

Appendix Number 4

to

Final Report

Project No. 2-0711

Grant No. OEG-0-72-0711

New Statistical Techniques to Evaluate Criterion-Referenced Tests Used in Individually Prescribed Instruction

Melvin R. Novick
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

December 21, 1973

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

This article, "High School Attainment: An Example of a Computer-Assisted Bayesian Approach to Data Analysis" has been removed from this document due to copyright restrictions. The entire article appears in the journal International Statistical Review, Vol. 41, No. 2, August 1973.

ACT TECHNICAL BULLETIN NO. 11

INTERDIALECT TRANSLATABILITY OF THE BASIC PROGRAMMING LANGUAGE

by

Gerald L. Isaacs

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

March, 1973

Interdialect Translatability of the

BASIC Programming Language

by

Gerald L. Isaacs
The American College Testing Program

Introduction

The BASIC (Beginner's All-purpose Symbolic Instruction Code) programming language is a mathematically-based conversational problem-solving language. It has wide application in business, scientific, and educational environments. It is powerful, efficient, flexible, and has the precision necessary for most tasks. Also, its syntax is simple and easy to learn. The BASIC programming language is simple enough so that an inexperienced programmer can use it and has enough power and flexibility so that the experienced programmer can write his programs efficiently. BASIC was first developed under Professors John G. Kemeny and Thomas E. Kurtz at Dartmouth College in 1963-1964. Since then, BASIC has been transformed into more than forty different major dialects. Each of these transformations has added to or modified the original language.

Due to the many differences among dialects of BASIC, unless care is taken in the initial programming it is both time consuming and difficult to readily translate a program from one dialect to another. However, if a few rules are followed, it may be possible to translate within a large set of dialects with a minimum of effort. In this paper we investigate this possiblity in some detail.

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

The specific purpose of this study is to identify and investigate in detail those BASIC dialects that would form a set in which translatability would be high if reasonable programming restrictions are imposed. Only BASIC dialects that are interactive have been studied. The following BASIC dialects are examined in this study.

BASIC FOUR BUSINESS BASIC

BURROUGHS 2500 BASIC

BURROUGHS 5500 BASIC

BURROUGHS 3500 BASIC

BURROUGHS 6700 BASIC (University of California San Diego)

CDC 6600 KRONOS/BASIC

CDC 6600 SCOPE/BASIC

COM-SHARE BASIC

COM-SHARE NEWBASIC

DARTMOUTH BASIC (sixth version)

DATA GENERAL EXTENDED BASIC

DEC PDP/8 BASIC (EDUSYSTEM 25 and 50)

DEC PDP/10 BASIC

DEC PDP/11 BASIC

GE MARK I BASIC

GE MARK II and GE MARK III BASIC

GE 255 TIME-SHARING BASIC

GENERAL AUTOMATION ADVANCED BASIC-16

HONEYWELL 200 BASIC

HONEYWELL 400 XBASIC

HONEYWELL 316, 516, and 716 BASIC

HONEYWELL 600 BASIC

HONEYWELL 1640 XBASIC

HP 2000B BASIC

HP 2000C BASIC

HP 2000E BASIC

HP 2000F BASIC

HP 3000 BASIC

IBM CPS/BASIC (University of Iowa)

IBM ITF/BASIC

IBM CALL/360-OS BASIC

IBM S3 MOD 6 BASIC

LEASCO BASIC

MICRODATA BASIC

MULTICOMP BASICX

NCR CENTURY 100 BASIC 1

NCR CENTURY 200 BASIC

Q-DATA BASIC 1

UNICOMP/COMP 16 or COMP 18 BASIC

UNIVAC 1100 UBASIC

UNIVAC 1100 UBASIC VERSION 2.0 (Mankato State)

UNIVAC 1100 (University of Maryland Release 1.3)

VARIAN 620 or V73 BASIC

WANG 2200 BASIC

WANG 3300 BASIC

WESTINGHOUSE BASIC II

WESTINGHOUSE BASIC III

XEROX BASIC

Due to the complexity and needs of our applications, we are mainly interested at this time in a multiuser system supporting a form of mass storage. Therefore, many single user systems or small systems such as the

PDP-8 EDUSYSTEM 5, 10, 20 have not been presented. However, most of the techniques discussed here also apply to many of the smaller systems.

Some of the dialects are immediately transportable from one computer in a manufacturer's line to another, e.g., the XEROX BASIC runs on the Sigma 5, 6, 7, 8, and 9. Also, several BASIC dialects are upward compatible on computers in the same line, e.g., the BASIC dialect on the Hewlett Packard 2000B will run on the 2000C, 2000E, and 2000F. For most dialects, some translation must be done if a program written in the BASIC of one computer is to be run on a second computer. This study was motivated by the desire to produce readily translatable conversational language interactive programs for computer-assisted data analysis and decision making in an educational environment. The conclusions of the study will, however, apply quite generally since the aforementioned applications are very demanding in terms of text handling capability, computational power, and formatting.

Important Programming Capabilities

There are four programming capabilities that should be present if a project of any magnitude or complexity is to be undertaken. The first of these is computational ability and precision. Of the more than forty-five dialects examined, all were found to provide at least six digits of accuracy and to support the basic arithmetic operations plus exponentiation. Some dialects provided accuracy of up to 15 or 16 digits. Obviously, dialects with only six digit accuracy will not be useful in many scientific applications. Also, there was a large variance as to the largest and smallest absolute number allowed. The smallest maximum absolute number was approximately 10^{37} while the largest minimum number was approximately 10^{-37} except for the Westinghouse and General Automation

BASICS which allowed approximately 10^{19} and 10^{-20}. For a translatable system, the questions of accuracy and precision will need to be considered carefully. A system can only be translated to dialects that provide the needed accuracy.

A second necessary capability of any dialect is that it has the ability to execute a program of the desired size. This may be accomplished in several ways. One method involves mass partition size. That is, a user is allowed as large a partition as is necessary for his task and is swapped in and out of core with many other users. This method may substantially add to cost and execution time. Further, when this method is used, a system that uses a monitor to sequentially execute several programs is not very feasible, since all the programs and the monitor must remain in core. In these circumstances, the user would load and execute each program independently. Such a procedure results in a tolerable inconvenience.

A second method that is used by many dialects is program chaining. This method allows the user to fit a very large program into a small partition by dividing the program into small segments and executing them separately in logical succession. There are two kinds of program chaining. The first calls for a complete overlay of the program in core, and the second, a chaining in which the user may specify where the overlay may begin.

The third method for accomplishing the execution of a large program is through the use of external subroutine calls. In this procedure, the user calls a subroutine that is maintained as a separate file. After it is executed, its core is released thus allowing additional portions of the program to be called into core without destroying existing code. There are some BASIC dialects such as IBM-CPS-BASIC, NCR-CENTURY-100-BASIC 1,

MICRODATA-BASIC, Q-DATA-BASIC-1, UNICOMP-BASIC, VARIAN-BASIC, and WESTINGHOUSE-BASIC-II which only allow a fixed area of core and do not permit the user any of the above options for increasing the size of the program to be executed. These dialects are inadequate for most complicated systems.

The third capability that a BASIC dialect should have is the means for accessing and creating external data files. Three levels of file capability are supported by the various BASIC dialects. One group of dialects offer no data file support, e.g., IBM-CPS-BASIC (UNIV of IOWA), WESTINGHOUSE-BASIC-II and III, GENERAL AUTOMATION-BASIC, HONEYWELL-316, 516, and 716 BASICS, MICRODATA-BASIC, Q-DATA-BASIC, BURROUGHS-2500 and 3500-BASICS, UNICOMP-BASIC, VARIAN-BASIC, COM-SHARE-BASIC, UNIVAC-1100-BASIC (UNIVERSITY OF MARYLAND V. 1.3), and NCR-CENTURY-100-BASIC-1. Presently, the NCR-CENTURY-200-BASIC has no file capability although it is promised in the near future. A second group of dialects supports only sequentially accessed data files. The latter group includes IMB-ITF-BASIC, IBM-S3-MOD-6-BASIC, GE-255-TIME-SHARING-EXTENDED-BASIC, HONEYWELL-1640-BASIC, WANG-3300 and 2200-BASICS, IBM-CALL/OS-BASIC, CDC-6000 KRONOS-BASIC, UCSD-B6700-BASIC, CDC-6000-SCOPE-BASIC, BURROUGHS-B5500-BASIC, DEC-PDP-8-BASIC (EDUSYSTEM 25 and 50), and UNIVAC-1100-UBASIC (MANKATO STATE VERSION 2.0). A third group of dialects supports both sequential access and random access files. Members of this group are HP2000F-2000E-2000C-2000B-BASIC, UNIVAC-1100-UBASIC, HP3000-BASIC, MULTICOMP-BASICX, BASIC-FOUR-BUSINESS-BASIC, XDS-BASIC, GE-MARK-I, MARK-II, AND MARK-III-BASICS, LEASCO-RESPONSE-I-BASIC, DARTMOUTH-BASIC, COM-SHARE-BASIC and NEWBASIC, DEC-PDP-10 and PDP-11-BASIC, HONEYWELL-200, 400, and 600-BASICS, and DATA-GENERAL-BASIC. The urgency of the need for random access files varies with the application. However, since

some type of file support is needed for nearly all applications, a minimum of sequential access to files is almost a must.

External files are used to store data that are too complicated and time consuming to recompute every time they are needed. Files also are needed to pass data between chained segments of a system; if the whole partition is overlayed. Also, files can be used to store results of computations so that the user may decrease the size of his program. In view of this, the BASIC dialects mentioned in the second and third groups of the previous paragraph are more adequate than the dialects in the first group.

A fourth important capability for a BASIC dialect is its conduciveness to generating formatted output. This is accomplished by means of the PRINT USING statement. This statement allows the user to determine what his output is going to look like. He may specify the number of digits to be outputted, the mode of output, and the column(s) in which the output is to appear. Also, the user may specify carriage control, e.g., number of spaces between lines. Most of these may also be accomplished using a PRINT statement. This is much less efficient, requires more programming, and cannot be accomplished in the case of specifying the number of digits. The PRINT USING statement has different syntax in almost every dialect. Therefore, it should be noted that if the PRINT USING is used, it must be modified when translating from one dialect to another. In some dialects the format to be followed is specified in the PRINT USING statement itself, while in others the format is in an IMAGE, FIELD, or format statement. Some dialects use Fortran format for output, e.g., MULTICOMP-BASICX. Others use an example output line with special characters denoting numeric output.

Not all systems have formatted output, e.g., HP2000B-BASIC, HP2000F BASIC, NCR-CENTURY-200-BASIC, CDC-6000-SCOPE and KRONOS-BASICS, and UNIVAC-1100-BASIC (UNIVERSITY OF MARYLAND). It is felt that a system should have a capability for formatted output. However, if it does not, the PRINT statement can provide many of the features of the PRINT USING command. Although the results may not be usually as appealing as with the PRINT USING statement, they provide a satisfactory alternative.

The translation of most statements in a BASIC dialect will be trivial or no translation will be necessary. Operands, relations, names, strings, arrays, functions, input, and branching can be translated with little effort or time. The three difficulties that will be encountered are file handling, chaining or subroutine calling, and output formatting. Since there is no exact standard for these areas, a knowledge of the statement formats in these areas can help to minimize the expenditure of time and energy.

Comparison of Elements

Operations and Relations:

All BASIC dialects use the same symbols for addition +, subtraction -, multiplication *, and division /. However, there is no standard operator for exponentiation. Different dialects use the following symbols: **, ↑, ^ . The most frequent symbol used for exponentiation is ↑ . If exponentiation can be avoided, translatability in operands is achieved. The string operation of concatenation is not implemented on all dialects. For those in which it is implemented, ampersand (&), plus (+), comma (,), STR, or CAT$ are used. A few of the dialects such

as BASIC-FOUR, VARIAN, COM-SHARE-NEWBASIC, HP2000B, HP2000C, HP2000E, HP2000F, HP3000, PDP-11, LEASCO, WANG-3300, WESTINGHOUSE, GENERAL-AUTOMATION, and the UNIVAC-1100 implement the logical operands of AND, OR, and NOT. The PDP-11, COM-SHARE-NEWBASIC, and UNIVAC-1100 BASIC dialects also support logical equivalence (EQV, EQU, and EQU, respectively), exclusive or (EOR, XOR, and XOR, respectively), and implication (IMP). GENERAL-AUTOMATION also supports exclusive or (XOR). The logical relations symbols for less than (<), greater than (>), not equal (<>), less than or equal (<=), and equal (=) are standard across all the BASIC dialects except for the UNIVAC-1100-UBASIC (VERSION 2.0 MANKATO STATE COLLEGE) dialect which uses LSS for less than, GRT for greater than, NEQ for not equal, LEQ for less than or equal, and EQU for equal, and MICRODATA which uses # for not equal. The logical relation greater than or equal (>=) is standard across all BASIC dialects except for the UNIVAC-1100 (VERSION 2.0 MANKATO STATE COLLEGE) and HONEYWELL-200-BASIC dialects which use the symbols GEQ and =>, respectively.

Names:

In the BASIC programming language, there can be up to five types of variable names. These are array variable names, numeric variable names, string variable names, integer variable names, and user defined function names. A numeric variable name should be either a letter or a letter followed by a single digit. While the IBM-BASIC dialects allow the special characters of $, @, and # to be used anywhere a letter may be used, and IBM-CPS-BASIC allows a single letter or a letter followed by another letter or a number, for reasons of translatability these conventions should not be used. String variables are used in

all BASIC dialects except NCR-CENTURY-100-BASIC-1, BURROUGHS-2500 and 3500-BASICS, UNICOMP-BASIC, VARIAN-BASIC, WESTINGHOUSE-BASIC-II, HONEYWELL-316, 516, and 716-BASICS, MICRODATA-BASIC, and Q-DATA BASIC-1.

There are two conventions used for string variable names. The first is a letter followed by a $. The second is a numeric name followed by a $. For translatability the first convention, a letter followed by a $, should be used. Integer variable names are only allowed in the PDP-11-BASIC and HP3000-BASIC and should be avoided. Array variable names should be confined to a single letter that has not been used elsewhere. Some dialects allow any numeric name to be an array name and allow the same name to be both an array variable name and a numeric variable name. In the interest of translatability, array variable names should be confined to a single unique letter. User defined function names are standard in all BASIC dialects except the NCR-200-BASIC, UNICOMP-BASIC and PDP-11-BASIC. There are no user defined functions in the NCR-200 and UNICOMP-BASIC dialects. The PDP-11-BASIC allows the user defined function to be FN followed by any numeric variable name. All other BASIC dialects limit a user defined function to FN followed by a single letter. The general convention of FN letter should be used.

Strings:

All BASIC dialects for the UNICOMP-BASIC, BURROUGHS-2500 and 3500-BASICS, VARIAN-BASIC, WESTINGHOUSE-BASIC-II, HONEYWELL-316, 516, and 716-BASICS, MICRODATA-BASIC, Q-DATA-BASIC-1, and NCR-CENTURY-100-BASIC 1 have string handling capabilities. However, these dialects still allow strings in PRINT statements. String constants are enclosed in quotes.

In all BASIC dialects except IBM-CPS-BASIC, IBM-S3-MOD-6-BASIC, UNIVAC-1100-UBASIC (VERSION 2.0 MANKATO STATE), and XDS-BASIC, double quotes (") may be used. In the exceptions, single quotes (') are used. Therefore, if translation is to take place between dialects that use the different types of string quotes, a user must be sure to change all the quotes. Strings vary in length in the BASIC dialects. The shortest string length is 6 characters and the longest string length is over 32,000 characters. There are two groups of dialects, those that allow a maximum of 6 to 22 characters and those that allow string length greater than or equal to 72. The dialects that provide a string length less than or equal to 22 characters are DEC-PDP-8-BASIC (EDUSYSTEM 25 and 50), BURROUGHS-B5500-BASIC, IBM-CPS-BASIC, IBM-S3-MOD-6-BASIC, WANG-3300-BASIC, IBM-ITF-BASIC, XDS-BASIC, IBM-CALL/360-OS-BASIC, HONEYWELL-200-BASIC, GE-255-TIME-SHARING-BASIC, GE-MARK-I-BASIC, NCR-CENTURY-200-BASIC, and UCSD-B6700-BASIC. Several of the BASIC dialects provide string processing functions from which substrings, positions, lengths, and other data may be obtained. It should be noted that these functions are not translatable and should not be used if the system is to be translated. If string handling is not needed, then all BASIC dialects can be considered. But if a long string (greater than 22) is needed, then translatability is limited.

Arrays:

All BASIC dialects allow use of arrays to store data. An array may have, at most, two dimensions in all BASIC dialects except CDC-6600-SCOPE-BASIC, BASIC-FOUR-BUSINESS-BASIC, VARIAN-BASIC, CDC-6600-KRONOS-BASIC and the HONEYWELL-200-BASIC, which allow three dimensions, UNIVAC-1100-BASIC which allows four dimensions, and COM-SHARE-NEWBASIC,

WESTINGHOUSE-BASIC, GENERAL-AUTOMATION-BASIC, and HONEYWELL-400, 316, 516, and 716-BASICS which allow as many dimensions as will fit in one statement. All BASIC dialects have some limit on the number of elements. In IBM-CPS-BASIC the limit is 500 elements per array. But in most BASIC dialects it is limited only by the amount of core that is available. Arrays that do not appear in a dimension (DIM) statement are dimensioned ten, or ten by ten, or ten by ten by ten depending upon use and system, in all dialects except PDP-8-BASIC (EDUSYSTEM 25 and 50), BASIC-FOUR-BUSINESS-BASIC, GENERAL-AUTOMATION-BASIC, WANG-3300 and 2200-BASICS, and NCR-CENTURY-200-BASIC. Therefore, all arrays should be dimensioned for translatability. Depending upon the dialect, arrays start at zero or one. But in matrix (MAT) operations, the zero elements are ignored anyway. All BASIC dialects have the MAT operations addition, subtraction, scaler multiplication, multiplication, transposition, and inversion except the PDP-8-BASIC (EDUSYSTEM 25 and 50), NCR-CENTURY-200-BASIC, NCR-CENTURY-100-BASIC-1, BASIC-FOUR-BUSINESS-BASIC, UNICOMP-BASIC, WESTINGHOUSE-BASIC-II, HONEYWELL-316, 516, and 716-BASICS, MICRODATA-BASIC, Q-DATA-BASIC, WANG-2200-BASIC, and UCSD-B6700-BASIC which do not support MAT operations. Also, there is an identity matrix (IDN), a matrix of all ones (CON) and a zero matrix (ZER) in all dialects that have the MAT commands. All dialects that support the MAT commands also support a form of matrix input and output. In addition, some support a file input and output for matrices. Whether an array is translatable or not depends upon several factors, including program size and partition size. The PDP-11-BASIC, HONEYWELL-400-BASIC, and COM-SHARE-BASIC allow arrays to reside on disc in what is called their virtual storage. But these are the only dialects that support a feature like this.

Complex Variables:

Only the HP3000-BASIC and COM-SHARE-NEWBASIC dialects allow the use of complex variables. Therefore, this capability should be avoided.

Functions:

BASIC functions are divided into two types. The first type includes all functions permanently resident in the system. All BASIC dialects support the following system functions:

ABS	Absolute value (except UNICOMP-BASIC)
ATN	Arctangent (except WESTINGHOUSE-BASIC-II and BASIC-FOUR; and BURROUGHS-2500 and 3500 and COM-SHARE-NEWBASIC which use ATAN)
COS	Cosine (except WESTINGHOUSE-BASIC-II and BASIC-FOUR)
EXP	Exponentiation (except BASIC-FOUR-BASIC)
INT	Largest integer (except UNICOMP-BASIC)
LOG	Common logarithm (except BASIC-FOUR-BASIC)
RND	Randomization (except WESTINGHOUSE-BASIC-II, UNICOMP-BASIC, and BASIC-FOUR-BASIC; and COM-SHARE-NEWBASIC which uses num)
SGN	Sign (except UNICOMP-BASIC
SIN	Sine (except WESTINGHOUSE-BASIC-II and BASIC-FOUR-BUSINESS-BASIC)
SQR	Square root (except BASIC-FOUR-BUSINESS-BASIC and WESTINGHOUSE-BASIC-II)
TAN	Tangent [except IBM-CPS-BASIC (UNIVERSITY of IOWA), UNICOMP-BASIC, BASIC-FOUR-BUSINESS-BASIC, and WESTINGHOUSE-BASIC-II]

The preceding system functions can be used freely unless in one of the exception dialects. The various dialects also support many other functions that should be avoided.

The second type of function is a user defined function. These functions pass one or several arguments depending on the dialect. Also, some dialects allow multiple line definitions. To be truly translatable, only single line definitions that pass at most one variable should be used. All BASIC dialects allow user defined functions except for NCR-CENTURY-200-BASIC and UNICOMP-BASIC.

Branching:

There are four types of statements used in BASIC for branching purposes. The first type of branching statement is the FOR statement. This loops control between the FOR statement and its corresponding NEXT statement until a counter reaches a limit. The format that is used in all BASIC dialects is:

```
FOR variable = initial value TO limit STEP increment.
NEXT variable
```

Initial value, increment and limit may be any expression in all BASIC dialects except BASIC-FOUR-BUSINESS-BASIC, and COM-SHARE-BASIC. In COM-SHARE-BASIC limit must be a number and in BASIC-FOUR-BUSINESS-BASIC initial value, increment and limit may be variables. In all BASIC dialects the loop works in the following manner:

1) The variable is set equal to the initial value.
2) Test if variable is searched or passed the limit.
    a) Execute loop if limit has not been reached.
    b) Exit loop if limit has been reached.
3) Add increment to variable.
4) Go back to step 2.

In all BASIC dialects loops may be nested, but maximum nesting permitted varies between dialects. If the user picks five as the deepest loops can be nested, then the system should be translatable.

The second type of branching statement is the IF statement. There are many forms of the IF statement in the BASIC dialects; but there is one that holds across all dialects. That is:

IF expression logical operator expression THEN line number..

The third type of branching statement is the GOTO statement. There are two forms of this statement, the simple GOTO and the computed GOTO. The computed GOTO is not implemented in all dialects and should be avoided. The simple GOTO is standard in all dialects as:

GOTO line number.

The word GOTO may also be GO TO in some dialects but it is not clear from the manuals which is accepted.

The fourth type of branching statement is the GOSUB statement. Here there are also two forms, the simple GOSUB and the computed GOSUB. The computed GOSUB is not universal and should be avoided. The simple GOSUB has the following syntax

GOSUB line number.

This form is standard across all BASIC dialects.

Therefore, if the preceding forms of the branching statements are used, the users' system will be translatable in terms of branching.

Input:

In the BASIC programming dialect there are two methods for accepting input. The first method is the READ-DATA statement pair. These two statements are completely translatable across all BASIC dialects except BASIC-FOUR-BUSINESS-BASIC which does not allow READ-DATA pairs. The form of these two statements is:

READ var 1, var 2, ... var n

DATA constant, constant, ... constant.

The only restriction is that in the NCR-CENTURY-100-BASIC-1, BURROUGHS-2500 and 3500-BASICS, COM-SHARE-BASIC, WESTINGHOUSE-BASIC-II, HONEYWELL-315, 516, and 716-BASICS, MICRODATA-BASIC, and Q-DATA-BASIC do not allow string variables or constants in the READ or DATA statements. The next read position in the data list can be reset to the beginning using the RESTORE command in all BASIC dialects except DARTMOUTH-BASIC which uses the RESET statement and UNICOMP-BASIC which has no prevision for starting over in a DATA statement.

The second method for accepting input is via the INPUT statement. In BASIC the INPUT statement accepts input from the user's terminal. The INPUT statement has the following syntax:

INPUT var 1, var 2, ... var n.

This syntax is constant over all BASIC dialects for this statement, although the same dialects that do not allow strings in READ-DATA pairs do not allow strings here. Thus, these statements are easily translatable.

Files:

The least translatable of all the statements are the file handling statements. Different dialects have different methods for handling files. In some dialects the user allocates a file name with a FILE statement, a FILES statement or an ASSIGN statement depending upon the dialect. Other dialects implicitly do this in the OPEN statement or first access. Backspacing and rewinding of files are allowed in a few dialects. Some dialects read from files with an INPUT statement while others use a READ statement. Also, PRINT and WRITE statements are used for writing into files in different dialects. Some dialects sense for end of file

with an IF END statement, others use a NODATA statement, while others use an ENDFILE statement. File names are determined from dialect to dialect and even from installation to installation within a dialect. Therefore, file handling is not directly translatable and the program writer should attend carefully to file input and file output statements when designing translatable programs.

Miscellaneous:

There are several aspects of BASIC that do not fall into any of the above categories. The first of these is the range on line numbers across the different dialects. The maximum range found was from 0 to 99999999. However, all dialects except IBM-CPS-BASIC and the PDP-8-BASIC accept line numbers from 1 to 9999. IBM-CPS-BASIC has a range from 1 to 999 and PDP-8-BASIC (EDUSYSTEM 25 and 50) has a range from 1 to 2046. Therefore, one should use line numbers only from 1 to 9999 for translatability. Unless either of the two above exceptions are to be used.

Another feature is comments or remarks; these can be fully translatable if the syntax is:

REM message.

Some dialects zero all variables before they are used, but this should not be taken for granted across all the dialects.

Also, certain dialects such as PDP-11-BASIC, COM-SHARE-NEWBASIC, HONEYWELL-316, 516, 716, and 600-BASICS, HP3000-BASIC, WANG-3300, and 2200-BASICS, and the HONEYWELL-200-BASIC allow multiple statements on a single line. This feature should not be used.

The keyword LET should not be dropped from assignment statements since many of the dialects require it. Also, only one variable should be assigned at a time even though many dialects allow multiple assignments. The format appears as:

LET var = expression.

The following three statements:

STOP

END

RETURN

are completely translatable when used in the above syntax. Some dialects allow a comment to follow. This should be avoided for reasons of translatability.

Summary of Rules for Translatability

1) Avoid the use of exponentiation if possible or use ↑ in all dialects where it is permitted.
2) Do not use logical arithmetic (OR, AND, NOT, etc.).
3) Use the following logical relations: <, >, <>, =, <=, and >= whenever permitted.
4) Use a single letter or a letter followed by a number for a numeric variable name.
5) Use a single letter followed by a $ for string variable names.
6) Use a unique letter for an array variable name.
7) Use FN followed by a single letter for a user defined function name.
8) Use double quotes (") whenever possible.
9) Decide on what length strings are going to be allowed and translate your system within the group your string length specifies.
10) Avoid string handling system functions.
11) Use at most two dimension arrays.
12) Start arrays at 1.
13) Take advantage of the MAT command where applicable.

14) Only the system functions listed should be used.

15) Use only single line user defined functions.

16) Nest loops at most five deep.

17) Limit the following statements to the listed format.

FOR variable = variable TO expression STEP expression

NEXT variable

IF expression-operator-expression THEN line number

GOTO line #

GOSUB line #

READ var 1, ...

DATA constant 1, ...

INPUT var 1, ...

STOP

END

RETURN

RESTORE

REM message

LET var = expression

18) Do not use multiple statements on a single line.

19) Line numbers should run from 1 to 9999.

20) Do not expect the system to zero all variables.

21) Avoid integer and complex variables.

Translatable BASIC Dialects

Most of the problems in translating one dialect to another are a matter of changing a keyword or format. These changes can be made to the whole program at one time using the edit features of the system. There are two features that must be changed or at least checked very

closely. These are the file handling and formatted output capabilities. These are not difficult changes to make, but must be considered carefully.

It was found that the dialects studied fell into three categories. The first of these categories contains those dialects that are missing a critical element. These are:

BURROUGHS-2500-BASIC (no data file capability, no capacity for chaining etc.)

BURROUGHS-3500-BASIC (no data file capability, no capacity for chaining etc.)

COM-SHARE-BASIC (manual chaining only)

GENERAL-AUTOMATION-BASIC (no data file capability)

HONEYWELL-316, 516, and 716-BASICS (chain only FORTRAN or ASSEMBLER routines)

IBM-CPS-BASIC (UNIV OF IOWA) (no data file capability, no capacity for chaining etc.)

MICRODATA-BASIC (no data file capability, no capacity for chaining etc.)

NCR-CENTURY-100-BASIC (no data file capability, no capacity for chaining etc.)

NCR-CENTURY-200-BASIC (no data file capability at this time)

Q-DATA-BASIC (no data file capability, no capacity for chaining etc.)

UNICOMP-BASIC (no data file capability, no capacity for chaining etc.)

UNIVAC-1100 (UNIV OF MARYLAND VERSION 1.3) (no data file capability)

VARIAN-BASIC (no data file capability)

WESTINGHOUSE-BASIC-II (no data file capability, no capacity for chaining etc.)

WESTINGHOUSE-BASIC-III (no data file capability)

The second category contains those dialects that only do not have formatted output capability.

BASIC-FOUR-BUSINESS-BASIC

BURROUGHS-5500-BASIC

DEC-PDP-8-BASIC (EDUSYSTEM 25 and 50)

GE-255-TIME-SHARING-BASIC

HP2000B-BASIC

HP2000E-BASIC

UCSD-B6700-BASIC (UNIVERSITY OF CALIFORNIA, SAN DIEGO)

UNIVAC-1100-UBASIC (VERSION 2.0 MANKATO STATE COLLEGE)

Also included in this category are those dialects that issue mass storage in place of chaining or external subroutine capability.

CDC-6600-KRONOS-BASIC (also no formatted output)

CDC-6600-SCOPE-BASIC (also no formatted output)

IBM-CALL/OS-360-BASIC

IBM-ITF-BASIC

The third category contains those dialects which are preferred.

COM-SHARE-NEWBASIC

DATA-GENERAL-BASIC

DARTMOUTH-BASIC

DEC-PDP-10 BASIC

DEC-PDP-11 BASIC

GE-MARK-I-BASIC

GE-MARK-II-BASIC

GE-MARK-III-BASIC

HONEYWELL-200-BASIC

HONEYWELL-400-XBASIC

HONEYWELL-600-BASIC

HP2000C-BASIC

HP2000F-BASIC

HP3000-BASIC

IBM-S3-MOD-6-BASIC

LEASCO-BASIC

HONEYWELL-1640-BASIC

MULTICOMP-BASICX (UNIV OF MASS, AMHERST, CDC-3600)

WANG-3300-BASIC

UNIVAC-1100-UBASIC

XDX-BASIC

WANG-2200-BASIC

Therefore, following the recommended translatability rules, a user should be able to obtain a system tnat is translatable with a minimum of effort and time within the third catagory and translatable with greater difficulty and expense in the second category. It should be noted that a program usually runs at a slower speed on a small machine than on a large machine.

The information provided above is a synopsis of extensive charts comparing the above dialects. These charts are available from the author. All information was obtained from manufacturers' manuals and is subject to change. It can clearly be seen that BASIC translatability is a fact and can be performed easily if a few rules are followed.

References

BASIC/Four Corporation
Reference Manual
Business BASIC
October 3, 1972 .

Burroughs
B5500
Time Sharing System
BASIC Reference Manual
November 15, 1969, #AA22366

Burroughs
Medium Systems
BASIC and CANDE
Reference Manual
1971, #1054905

Computer Center
University of California
San Diego
BASIC User's Manual
February 1972

Computer Science Center
University of Maryland
BASIC for the UNIVAC 1108
September 1968, #800002

Com-Share
BASIC Reference Manual
May 1970, #9005-5

Com-Share
NEWBASIC
Reference Manual
February 1971

Control Data
6000 Series
Computer Systems
KRONOS 2.0
BASIC Reference Manual
January 27, 1971, #59150800

Control Data
Cyber 70 Computer Systems
BASIC Language Reference Manual
February 29, 1972, #60306200

Data General Corporation
Extended BASIC
User's Manual
1972, #093-000065-01

Digital Equipment Corporation
Programming Languages
PDP-8 Handbook Series
May 1970

Digital Equipment Corporation
BASIC-PLUS
Language Manual
1972, #DEC-11-ORBPA-A-D

Digital Equipment Corporation
DEC SYSTEM 10
Mathematical Languages Handbook
1972, #DEC-10-AFDO-D

Digital Equipment Corporation
Edusystem Handbook
Edusystem-25
January 1973

Digital Equipment Corporation
Edusystem Handbook
Edusystem-50
January 1973

GE Information Services
MARK II
BASIC Language
June 1972, #3200.01D

GE Information Systems
GE 255 Time-Sharing
BASIC Language Extended
August 1969, #CPB-1624

GE Information Systems
MARK I
BASIC Language
February 1970, #202026B

General Automation
BASIC-16
Reference Manual
October 1972, #88A00256A

Hewlett Packard
2000B:
A Guide to Time-Shared BASIC
August 1970, #HP 02000-90010

Hewlett Packard Corporation
2000C:
A Guide to Time-Shared BASIC
April 1971, #HP 02000-90016

Hewlett Packard
3000 BASIC
External Reference Specification
September 15, 1971

Hewlett Packard Corporation
2000E:
A Guide to Time-Shared BASIC
August 1972, #02000-90048

Honeywell
Series 200 BASIC SYSTEM
Language Manual
March 1970, #133.0005.1100.0-082

Honeywell
Series 1640
BASIC Language Manual
October 1970, #AD12

Honeywell
Series 400
TSPS BASIC Language
August 1971, #BR03

Honeywell
Series 600/6000
Time-Sharing BASIC
September 1971, #BR36, Rev. 1

Honeywell
Series 16
BASIC Language
March 1972, #AB85, Rev. 1

IBM
Call/360-OS
BASIC Language Reference Manual
January 1971, #GH20-0699-2

IBM System/360-OS(TSO)
ITF: BASIC
Terminal User's Guide
June 1, 1971, #SC28-6840-0

IBM System/3 Model 6
System/3 BASIC
Reference Manual
June 1972, #GC34-0001-2

Kiewit Computation Center
Dartmouth College
BASIC Sixth Edition
1971, #0-87451-063-5

Leasco
Response I-BASIC
Reference Manual
June 1971

Mankato State College
UBASIC VERSION 2.0
Programmers Reference Manual

Microdata
BASIC
User's Manual
1972

Multicomp
Reference Manual/Language
BASIC and BASICX
1970, #5M.0315.01

National Cash Register
Time-Sharing System
BASIC
General Information
November 1970, #ST-9462-20

Q-DATA Corporation
BASIC-1
November 1970

UNI-COMP
BASIC Manual
#P701110, 038

UNIVAC
1100 Series
BASIC
1972, #UP7925

University Computer Center
University of Iowa
CPS
1971

Varian
Software Handbook
1971

Wang
3300 BASIC
Reference Manual
1971

Wang
2200 BASIC
Preliminary Copy

Westinghouse
BASIC II & III
Reference Manual
March 1971, #25REF-012

Xerox Data Systems
Xerox BASIC
Sigma 5-9 Computers
Language and Operations
Reference Manual
March 1972, #90-15-46D

A C T   T E C H N I C A L   B U L L E T I N

S U P P L E M E N T   N O .   1 1 - 1

A TABULAR SURVEY OF BASIC COMPUTER SYSTEMS*

by

Gerald L. Isaacs

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

March, 1973

*The research reported herein was performed pursuant to Grant No. OEG-C-72-0711 with The Office of Education, U. S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

The purpose of this set of tables is to supplement the basic report on <u>Interdialect Translatability of the BASIC Programming Language</u> (Gerald L. Isaacs, Technical Bulletin No. 11, The American College Testing Program) and to give a quick, clear, concise, updated view of many of the BASIC languages as supported on different computer systems. As shown, each system has irs own set of commands and its own set of capabilities. Preceding the tables is a listing of the conventions used in the tables, a summary of file capabilities of the various systems and a list of references. Immediately following are some late arriving materials that could not conveniently be included in the tables. The summary presented here includes all information we have been able to gather as of March 9, 1973. We have given the authors of all dialects surveyed an opportunity to respond to a preliminary draft, and we have worked closely with those authors who have responded to requests. Nevertheless, we cannot believe that we have attained 100% accuracy and even if we did that accuracy would soon decay as a result of the continuing fast pace of improvement now evidencing itself. We should note specifically that we have <u>not</u> credited various dialects with features that are "promised for delivery in the near future" or even those which we are told exist but "are not yet documented". For this reason, we urge any potential user to check with the relevant manufacture before dismissing from consideration any system that seems attractive. At the same time, we urge manufacturers to supply us with documentation of improvements so that we can keep our charts up to date.

Gerald L. Isaacs
March 9, 1973

The following conventions are used in the tables.

--------	not available
num or n	number
var	variable
exp	expression
arg	argument
numlist	number list
val list	value list
var list	variable list
op	operator
str	string
param	parameter

File Capability

BURROUGHS-B2500	none
BURROUGHS-B3500	none
GENERAL AUTOMATION BASIC-16 ADVANCED	none
HONEYWELL 316, 516, and 716	none
IBM/CPS (Univ of Iowa)	none
MICRODATA	none.
NCR CENTURY 100	none
NCR CENTURY 200	no files as of now
Q-DATA BASIC-1	none
UNICOMP-COMP 16, COMP 18	none
UNIVAC 1100 (Univ of Maryland) Version 1.3	none
WESTINGHOUSE BASIC II	none
WESTINGHOUSE BASIC III	none
BURROUGHS-B5500	sequential
CDC 6000-KRONOS	sequential
CDC 6000-SCOPE	sequencial
DEC-PDP8 (Edusystem 25 and 50)	sequential
GE 255 Time Sharing Extended BASIC	sequential
HONEYWELL 1640	sequential
IBM/CALL/360-OS	sequential
IBM/ITF	sequential
IBM S3 MOD 6	sequential
UCSD-B6700	sequential
UNIVAC 1100-UBASIC (Mankato State) Version 2.0	sequential
WANG 3300	sequential
WANG 2200	sequential

BASIC 4-BUSINESS BASIC	sequential and random access
COM-SHARE BASIC	sequential and random access
COM-SHARE NEWBASIC	sequential and random access
DARTMOUTH	sequential and random access
DATA GENERAL	sequential and random access
DEC-PDP 10	sequential and random access
DEC-PDP 11	sequential and random access
GE MARK I	sequential and random access
GE MARK II	sequential and random access
GE MARK III	sequential and random access
HONEYWELL 200	sequential and random access
HONEYWELL 400	sequential and disc arrays
HONEYWELL 600	sequential and random access
HP2000B	sequential and random access
HP2000C	sequential and random access
HP2000E	sequential and random access
HP2000F	sequential and random access
HP3000	sequential and random access
LEASCO-RESPONSE I	sequential and random access
MULTICOMP BASICX	sequential and random access
UNIVAC 1100-UBASIC Version 3.2	sequential and random access
XDS-BASIC	sequential and random access

References

BASIC/Four Corporation
Reference Manual
Business BASIC
October 3, 1972

Burroughs
B5500
Time Sharing System
BASIC Reference Manual
November 15, 1969, #AA22366

Burroughs
Medium Systems
BASIC and CANDE
Reference Manual
1971, #1054905

Computer Center
University of California
San Diego
BASIC User's Manual
February 1972

Computer Science Center
University of Maryland
BASIC for the UNIVAC 1108
September 1968, #800002

Com-Share
BASIC Reference Manual
May 1970, #9005-5

Com-Share
NEWBASIC
Reference Manual
February 1971

Control Data
6000 Series
Computer Systems
KRONOS 2.0
BASIC Reference Manual
January 27, 1971, #59150800

Control Data
Cyber 70 Computer Systems
BASIC Language Reference Manual
February 29, 1972, #60306200

Data General Corporation
Extended BASIC
User's Manual
1972, #093-000065-01

Digital Equipment Corporation
Programming Languages
PDP-8 Handbook Series
May 1970

Digital Equipment Corporation
BASIC-PLUS
Language Manual
1972, #DEC-11-ORBPA-A-D

Digital Equipment Corporation
DEC SYSTEM 10
Mathematical Languages Handbook
1972, #DEC-10-AFDO-D

Digital Equipment Corporation
Edusystem-25
User's Manual

Digital Equipment Corporation
Edusystem-50
User's Manual

GE Information Services
MARK II
BASIC Language
June 1972, #3200.01D

GE Information Systems
GE 255 Time Sharing
BASIC Language Extended
August 1969, #CPB-1624

GE Information Systems
MARK I
BASIC Language
February 1970, #202[illegible]6B

General Automation
BASIC-16
Reference Manual
October 1972, #88A00256A

Hewlett Packard Corporation
2000B:
A Guide to Time-Shared BASIC
August 1970, #HP 02000-90010

Hewlett Packard Corporation
2000C:
A Guide to Time-Shared BASIC
April 1971, #HP 02000-90016

Hewlett Packard Corporation
3000 BASIC
External Reference Specification
September 15, 1971

Hewlett Packard Corporation
2000E:
A Guide to Time-Shared BASIC
August 1972, #02000-90048

Honeywell
Series 200 BASIC SYSTEM
Language Manual
March 1970, #133.0005.1100.0-082

Honeywell
Series 1640
BASIC Language Manual
October 1970, #AD12

Honeywell
Series 400
TSPS BASIC Language
August 1971, #BR03

Honeywell
Series 600/6000
Time Sharing BASIC
September 1971, #BR36, Rev. 1

Honeywell
Series 16
BASIC Language
March 1972, #AB85, Rev. 1

IBM
Call/360-OS
BASIC Language Reference Manual
January 1971, #GH20-0699-2

IBM System/360-OS(TSO)
ITF: BASIC
Terminal User's Guide
June 1, 1971, #SC28-6840-0

IBM System/3 Model 6
System/3 BASIC
Reference Manual
June 1972, #GC34-0001-2

Kiewit Computation Center
Dartmouth College
BASIC Sixth Edition
1971, #0-87451-063-5

Leasco
Response 1-BASIC
Reference Manual
June 1971

Mankato State College
UBASIC VERSION 2.0
Programmers Reference Manual

Microdata
BASIC
User's Manual
1972

Multicomp
Reference Manual/Language
BASIC and BASICX
1970, #5M.0315.01

National Cash Register
Time-Sharing System
BASIC
General Information
November 1970, #ST-9462-20

Q-DATA Corporation
BASIC-1
November 1970

UNI-COMP
BASIC Manual
#P701110, 038

UNIVAC
1100 Series
BASIC
1972, #UP7925

University Computer Center
University of Iowa
CPS
1971

Varian
Software Handbook
1971

Wang
3300 BASIC
Reference Manual
1971

Wang
2200 BASIC
Preliminary Copy

Westinghouse
BASIC II & III
Reference Manual
March 1971, #25REF-012

Xerox Data Systems
Xerox BASIC
Sigma 5-9 Computers
Language and Operations
Reference Manual
March 1972, #90-15-46D

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C	1A
ACCESS	--------	--------	--------	--------	--------	--------	--------	
ACCURACY**	14 DIGITS	14 DIGITS	6-8 DIGITS	6 DIGITS	9 DIGITS	6-7 DIGITS	6-7 DIGITS	
AND	--------	--------	--------	--------	--------	AND	AND	
APPROXIMATELY EQUAL	--------	--------	--------	--------	--------	--------	--------	
ARRAY NAME**	letter or letter num	letter or ¢	letter	letter	letter	letter	letter	
ARRAY STARTING**	1	0	0	0	0	1	1	
ASSIGN	--------	--------	see FILE	--------	--------	see FILES	ASSIGN name, num, var, mask	
BACKSPACE	--------	--------	--------	--------	BACKSPACE #I exp BACKSPACE$: exp	--------	--------	
CALL	--------	--------	CALL "NAME": arg1, arg2 call by ref call by value arg1	CALL num, list	CALL routine name not standard	--------	--------	
CHAIN	--------	--------	CHAIN name SYSTEM WITH file #num1, #num2, ...	CHAIN name	CHAIN name CHAIN name, Password	CHAIN name	CHAIN name, exp	

**Commands and elements that can be used.

..	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
ACCESS	--------	--------	--------	--------	--------	ACCESS num, code	--------
ACCURACY**	15 DIGITS	7 DIGITS	8 DIGITS	15 DIGITS, double precision	8 DIGITS	11 DIGITS	16 DIGITS
AND	--------	AND	--------	AND	AND(exp1, exp2)	--------	--------
APPROXIMATELY EQUAL	--------	--------	--------	= =	--------	--------	--------
ARRAY NAME**	letter	letter	letter	letter followed by num	letter	letter	letter
ARRAY STARTING**	1	1	0	0	0	0	1
ASSIGN	--------	ASSIGN (name, num, var, mask)	--------	see OPEN	--------	ASSIGN name to num	--------
BACKSPACE	--------	--------	--------	--------	--------	--------	--------
CALL	--------	--------	--------	--------	CALL name (parameter list)	--------	--------
CHAIN	--------	CHAIN name CHAIN name, line #	CHAIN name CHAIN name, exp	CHAIN name, line #	CHAIN name, n or CHAIN *name, n or CHAIN: name, n or CHAIN: *name, n	CHAIN name clears storage	CHAIN name; password: num

**Commands and elements that can be used.

1B

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F	1C
ACCESS	-----	-----	-----	-----	-----	-----	-----	
ACCURACY**	15 DIGITS	7 DIGITS	10 DIGITS	14 DIGITS	7 DIGITS	11 DIGITS	6 to 7 DIGITS	
AND	-----	-----	-----	-----	-----	-- --	AND	
APPROXIMATELY EQUAL	-----	-----	-----	-----	-----	-----	-----	
ARRAY NAME**	letter or $ or @ or #	letter	letter	letter or letter number	letter	letter	letter	
ARRAY STARTING**	1	0	0	1	1	0	1	
ASSIGN	-----	-----	-----	-----	-----	-----	ASSIGN name, num, name, mask	
BACKSPACE	-----	-----	BACKSPACE #exp	-----	-----	-----	-----	
CALL	-----	-----	CALL name CALL name num num=user #	-----	-----	-----	-----	
CHAIN	-----	-----	CHAIN name CHAIN name num num=user #	-----	CHAIN name	-----	CHAIN name, exp	

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC	1D
ACCESS	--------	--------	--------	--------	--------	-------	--------	
ACCURACY**	7 DIGITS	11 DIGITS	8 DIGITS	8 DIGITS	14 DIGITS	7 DIGITS	~ 7 DIGITS	
AND	--------	--------	--------	--------	AND	-------	AND	
APPROXIMATELY EQUAL	--------	--------	--------	--------	--------	-------	--------	
ARRAY NAME**	letter	num name	letter	letter	letter	letter	letter	
ARRAY STARTING**	0	1	1	1	0	0	1	
ASSIGN	--------	--------	--------	--------	--------	-------	--------	
BACKSPACE	--------	--------	--------	--------	--------	-------	--------	
CALL	--------	--------	--------	--------	--------	-------	CALL name, param, param1, ...	
CHAIN	--------	--------	--------	--------	RUN name	-------	--------	

**Commands and elements that can be used.

<table>
<tr><th></th><th>IBM S3
MOD 6
BASIC</th><th>GE 255
TIME SHARING
BASIC</th><th>COM-SHARE
BASIC</th><th>COM-SHARE
NEWBASIC</th><th>WESTINGHOUSE
BASIC II</th><th>WESTINGHOUSE
BASIC III</th><th>GENERAL AUTOMATION
ADVANCED
BASIC-16</th></tr>
<tr><td>ACCESS</td><td>--------</td><td>--------</td><td>--------</td><td>--------</td><td>--------</td><td>--------</td><td>--------</td></tr>
<tr><td>ACCURACY**</td><td>15 DIGITS</td><td>9 DIGITS</td><td>6 DIGITS</td><td>10 DIGITS
or
18 DIGITS</td><td>7-8 DIGITS</td><td>7-8 DIGITS</td><td>6 DIGITS</td></tr>
<tr><td>AND</td><td>--------</td><td>--------</td><td>--------</td><td>AND</td><td>AND</td><td>AND</td><td>AND</td></tr>
<tr><td>APPROXIMATELY EQUAL</td><td>--------</td><td>--------</td><td>--------</td><td>" #</td><td>--------</td><td>--------</td><td>--------</td></tr>
<tr><td>ARRAY NAME**</td><td>letter, $, #,
or @</td><td>letter</td><td>letter</td><td>letter</td><td>letter</td><td>letter</td><td>letter</td></tr>
<tr><td>ARRAY STARTING**</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>ASSIGN</td><td>ALLOCATE
name</td><td>see FILES</td><td>see OPEN</td><td>see OPEN</td><td>--------</td><td>--------</td><td>--------</td></tr>
<tr><td>BACKSPACE</td><td>--------</td><td>BACKSPACE # exp</td><td>--------</td><td>--------</td><td>--------</td><td>--------</td><td>--------</td></tr>
<tr><td>CALL</td><td>CALL name, line number operates like CHAIN</td><td>--------</td><td>--------</td><td>CALL name or
CALL FN letter
or
CALL $ name</td><td>--------</td><td>CALL (num, exp 1, exp 2, ...)</td><td>CALL name (exp 1, exp 2, ...)</td></tr>
<tr><td>CHAIN</td><td>see CALL</td><td>CHAIN name
or
CHAIN name,
line #</td><td>SCRATCH
PROCEED
Provides manual chaining</td><td>LINK "|name!"
LOAD "|name;"</td><td>--------</td><td>--------</td><td>--------</td></tr>
</table>

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3	1F
ACCESS	--------	--------	--------	--------	--------	--------	--------	
ACCURACY**	8 DIGITS	~ 6 DIGITS	~ 6 DIGITS	11 DIGITS	~ 8 DIGITS	6 to 7 DIGITS	8 DIGITS	
AND	AND(exp 1, exp 2)	--------	--------	--------	--------	AND	--------	
APPROXIMATELY EQUAL	--------	--------	--------	--------	--------	--------	--------	
ARRAY NAME**	letter	letter	letter	letter	letter	letter	letter	
ARRAY STARTING**	0	0	0	0	1	1	0	
ASSIGN	--------	--------	--------	--------	--------	see FILES	--------	
BACKSPACE	--------	--------	--------	BACKSPACE # num	BACKSPACE # num BACKSPACE : num	--------	--------	
CALL	CALL FUNC(exp 1, ... exp n)	CALL name	CALL(num, exp 1, ... exp n) Fortran or Assembler Only	CALL name or CALL name options	CALL name or CALL name, password	--------	CALL FUNC (exp 1, ... exp n)	
CHAIN	CHAIN name, num or CHAIN: or CHAIN* or CHAIN:*	--------	--------	CHAIN RUN: name options RUN may be any RUN command	CHAIN name, num CHAIN name CHAIN name, password, num	CHAIN name	--------	

**Commands and elements that can be used.

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200	
ACCESS	--------	--------	--------	--------	--------	--------	
ACCURACY**	9 DIGITS	6-7 DIGITS	6-7 DIGITS 11-12 DOUBLE PRECISION	8 DIGITS	9 DIGITS	13 DIGITS	
AND	--------	--------	AND	AND (exp 1, ... exp n)	--------	--------	
APPROXIMATELY EQUAL	--------	--------	--------	--------	--------	--------	
ARRAY NAME**	letter	letter	letter or letter digit	letter	letter	letter	
ARRAY STARTING**	0	0	1	1	0	1	
ASSIGN	--------	--------	ASSIGN name, exp, var, mask	--------	--------	SELECT options	
BACKSPACE	--------	--------	--------	--------	BACKSPACE # exp	--------	
CALL	--------	--------	CALL name	--------	CALL name	GOSUB' num (var 1, ... var n)	
CHAIN	--------	--------	CHAIN name, exp	CHAIN name CHAIN R name CHAIN num CHAIN R num	CHAIN name CHAIN name, num	LOAD name LOAD name, num	

**Commands and elements that can be used.

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C
CHANGE	--------	--------	CHANGE numlist to string BIT CHANGE string to numlist	--------	CHANGE numlist to string CHANGE string to numlist	--------	--------
CLOSE	--------	--------	--------	CLOSE FILE exp	--------	--------	--------
COMMON(FILE)	--------	--------	--------	--------	--------	--------	--------
COMMON(STORAGE)	--------	--------	--------	--------	--------	COM list	COM list
CONCATENATION	--------	--------	&	,	+	--------	--------
DATA**	DATA val list	DATA exp, ...	DATA val list	DATA val list	DATA exp, ...	DATA val list	DATA val list
DATA FILE	--------	--------	--------	--------	--------	--------	--------
DEF**	DEF FUNC (var) = exp one line	DATA FUNC (var list) = exp one line	DEF FUNC (var list) one or mult	DEF FUNC (var) = exp one line	DEF FUNC (var list) one or mult	DEF FUNC (var) = exp one line	DEF FUNC (var) = exp one line
DIM # (virtual stor)	--------	--------	--------	--------	--------	--------	--------
DIMENSION**	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions),... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10
END**	END	END	END	END	END	END	END
ENDFILE	--------	--------	--------	see IF END #	--------	see IF END	see IF END

**Commands and elements that can be used.

2B

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
CHANGE	--------	CONV	CHANGE string to num CHANGE num to string	CHANGE num to string CHANGE string to num	CHANGE numlist to string CHANGE string to numlist	CHANGE numlist to string CHANGE string to numlist	CHANGE string to num CHANGE num to string
CLOSE	CLOSE FILE 'name', ...	--------	--------	CLOSE list of exp	CLOSE num, ...	CLOSE #, name	CLOSE: num
COMMON(FILE)	--------	--------	--------	--------	--------	COMMON n	--------
COMMON(STORAGE)	--------	--------	--------	--------	--------	--------	--------
CONCATENATION	--------	--------	+	+	CAT$(str1, str2)	--------	+
DATA**	DATA val list	DATA val list	DATA val list	DATA val list	DATA val list	DATA val list DATA val:rep	DATA val list
DATA FILE	--------	--------	--------	--------	--------	DATAFILE = #, name	--------
DEF**	DEF FUNC (var list)=exp one line	DEF FUNC (var list)=exp one line	DEF FUNC (var list) one or mult line	DEF FUNC (var list) one line	DEF FUNC (var list) one or mult	DEF FUNC (var list)= exp one line	DEF FUNC (var list)=exp one line
DIM # (virtual stor)	--------	--------	--------	DIM # n name(dims)=exp	--------	--------	--------
DIMENSION**	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10
END**	END	END	END	END	END	END	END
ENDFILE	--------	--------	--------	--------	see ON	ENDFILE n	ENDFILE: exp, line #

**Commands and elements that can be used.

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
CHANGE	------	------	------	------	------	------	------
CLOSE	CLOSE expl,...	------	------	------	------	------	------
COMMON(FILE)	------	------	------	------	------	------	------
COMMON(STORAGE)	------	------	------	------	------	------	COM var1,...
CONCATENATION	------	------	------	------	------	&	------
DATA**	DATA val list	DATA val list no strings	DATA val list	DATA val list	DATA val list	DATA val list	DATA val list
DATA FILE	------	------	------	------	------	------	------
DEF**	DEF FUNC (var)=exp one line	DEF FUNC (var1,...)=exp one line	DEF FUNC (var1,...)=exp one line	DEF FUNC (var)=exp one line	------	DEF FUNC (var1,...)=exp one line	DEF FUNC (var)=exp one line
DIM # (virtual stor)	------	------	------	------	------	------	------
DIMENSION**	DIM name (dimensions), ... default 10	DIM name (dimensions), ...	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10	DIM name (dimensions), ...	DIM name (dimensions), ... default 10	DIM name (dimensions), ... default 10
END**	END comment	END	END	END	END	END	END
ENDFILE	------	------	see IF END	see NODATA	------	see IF END	see IF END

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC
CHANGE	--------	--------	--------	--------	--------	--------	--------
CLOSE	--------	--------	--------	--------	CLOSE (num)	--------	--------
COMMON(FILE)	--------	--------	--------	--------	--------	--------	--------
COMMON(STORAGE)	--------	--------	--------	--------	--------	--------	--------
CONCATENATION	--------	--------	--------	--------	--------	--------	--------
DATA**	DATA val list no strings	DATA val list	DATA val list no strings	DATA val list no strings	----------	DATA val list	DATA val list
DATA FILE	--------	--------	--------	--------	--------	--------	--------
DEF **	DEF FUNC (var)=exp one line	DEF FUNC (var1,...)=exp one line	DEF FUNC (var) = exp one line	DEF FUNC (var) = exp one line	DEF FUNC (var1,...)=exp one line	--------	DEF FUNC (var) = exp one line
DIM # (virtual stor)	--------	--------	--------	--------	--------	--------	--------
DIMENSION**	DIM var (dimensions), ... default 10	DIM var (dimensions), ... default 10	DIM var (dimensions), ... default 10	DIM var (dimensions), ... default 10	DIM var (dimensions), ...	DIM var (dimensions), ... default 10	DIM var (dimensions), ... default 30
END**	END	END	END	END	END	END	END
ENDFILE	--------	--------	--------	--------	see READ # see WRITE #	--------	--------

2D

**Commands and elements that can be used.

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16	2E
CHANGE	--------	--------	CHANGE str TO array var CHANGE array TO string var	CHANGE is a string function	--------	--------	--------	
CLOSE	CLOSE num or str var, ...	--------	CLOSE INPUT CLOSE OUTPUT	CLOSE exp	--------	--------	--------	
COMMON (FILE)	--------	--------	--------	--------	--------	--------	--------	
COMMON (STORAGE)	--------	--------	--------	--------	--------	--------	COM var, .	
CONCATENATION	LET STR (str var, num, num) = str var	--------	--------	--------	--------	+	--------	
DATA**	DATA val 1, val 2, ... val n	DATA val 1, val 2, ... val n	DATA val I, val 2, ... val n no strings	DATA val 1, ... val n	DATA val 1, val 2, ... val n no strings	DATA val 1, val 2, ... val n	DATA val 1, val 2, ... val n	
DATA FILE	--------	--------	--------	--------	--------	--------	--------	
DEF**	DEF FUNC(var)= exp single line	DEF FUNC(var)= exp single line	DEF FUNC(var)= exp single line	DEF FUNC(var 1, ...) = exp or DEF FUNC(var 1, ...) multiple line	DEF FUNC(var 1, ...) = exp single line	DEF FUNC(var 1, ...) = exp single line	DEF FUNC(var 1, ...) = exp single line	
DIM # (virtual stor)	--------	--------	DISC var(num), ... DISC var (num, num)...	--------	--------	--------	--------	
DIMENSION**	DIM var(num, num) ... default 10	DIM var(num, num), ... default 10	DIM var(num, num), ... default 10	DIM var(exp, ...), ... default 10	DIM var(n, ...), ... default n=10	DIM var(n, ...), ... default n=10	DIM var(n, ...), ...	
END**	END comment or END	END	END	END	END	END	END	
ENDFILE	--------	see IF END #	--------	ON ENDFILE GOTO line #	--------	--------	--------	

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
CHANGE	CHANGE string TO var CHANGE var TO string	--------	--------	CHANGE string TO var CHANGE var TO string	CHANGE string TO var CHANGE var TO string	--------	CHANGE string TO var CHANGE var TO string
CLOSE	CLOSE exp 1, ... exp n	--------	--------	CLOSE: name:	see SCRATCH	--------	--------
COMMON (FILE)	--------	--------	--------	--------	--------	--------	--------
COMMON (STORAGE)	--------	--------	--------	--------	--------	COM var 1, ... var n	--------
CONCATENATION	CAT$(str 1, str 2)	--------	--------	--------	--------	--------	--------
DATA**	DATA val 1, ..., val n	DATA val 1, ..., val n	DATA val 1, ..., val n no strings	DATA val 1, ..., val n	DATA val 1, ..., val n	DATA val 1, ..., val n	DATA val 1, ..., val n
DATA FILE	--------	--------	--------	--------	--------	--------	--------
DEF**	DEF FUNC(var), ..., var n)=exp DEF FUNC(var 1, ..., var n)	DEF FUNC(var)=exp single line	DEF FUNC(var)=exp single line	DEF FUNC(var)=exp single line	DEF FUNC(var)=exp or DEF FUNC(var), ..., var n)	DEF FUNC(var)=exp single line	DEF FUNC(var 1, ..., var n)=exp DEF FUNC(var 1, ..., var n)
DIM # (virtual stor)	--------	--------	--------	DISC var(num, num): name	--------	--------	--------
DIMENSION**	DIM var(num, num 1, num 2, num 3), ..., or ARRAYS to specify default default 10	DIM var(num, num), ... default 10	DIM var(num, ...), ... default 10	DIM var(num, ...), ... default 10	DIM var (num, num), ... default 10	DIM var(num, num), ... default 10	DIM var(num, num), ... default 10
END**	END	END	END	END	END	END	END
ENDFILE	ON ENDFILE exp GOTO line #	see IF END	--------	see IF END	see IF END	see IF END	--------

**Commands and elements that can be used.

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200	
CHANGE	--------	--------	CONVERT num TO string CONVERT string TO var	--------	--------	--------	
CLOSE	--------	--------	--------	FILEEND # num	--------	--------	
COMMON (FILE)	--------	--------	--------	--------	--------	--------	
COMMON (STORAGE)	--------	--------	COM list	COM var 1, ..., var n	--------	COM var 1, ..., var n	
CONCATENATION	--------	--------	+	--------	--------	--------	
DATA**	DATA val 1, val 2, ... val n no strings	DATA val 1, val 2, ... val n no strings	DATA val 1, ... val n	DATA val 1, ... val n	DATA val 1, ... val n	DATA val 1, ... val n	
DATA FILE	--------	--------	--------	--------	--------		
DEF**	DEF FUNC (var 1, var 2, ... var n) = exp single line	DEF FUNC (var) = exp single line	DEF FUNC (var 1, ... var n) = exp DEF FUNC (var 1, ... var n)	DEF FUNC (var) = exp single line	DEF FUNC (var) = exp single line	DEF FUNC (var) = exp single line	
DIM # (virtual stor)	--------	--------	--------	--------	--------	--------	
DIMENSION**	DIM var (dim, dim), ... default 10	DIM var (dim, dim), ... default 10	DIM var (dim, dim), ... default 10	DIM var (dim, dim), ...	DIM var (dim, dim), ... default 10	DIM var (dim, dim), ...	
END**	END	END	END	END	END	END	
ENDFILE	--------	--------	see IF END	see IF END	see IF END	--------	

**Commands and elements that can be used.

	BASIC 2.D CDC 66D0 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C
ENTER	--------	--------	--------	--------	--------	ENTER var, exp,... ENTER # var, exp, ...	ENTER var, exp,... ENTER # var, exp, ...
EQUAL **	=	=	=	=	= or .EQ.	=	=
EQUIVALENCE	--------	--------	--------	--------	--------	--------	--------
EXCHANGE	--------	--------	--------	--------	--------	--------	--------
EXCLUSIVE OR	--------	--------	--------	--------	--------	--------	--------
EXPONENTIATION **	↑ or **	**	^	↑	↑ or **	↑	↑
FIELD	--------	see IMAGE	see USING	--------	see USING	--------	see USING
FILE	--------	--------	FILE # exp: name	see OPEN	FILE # num, name FILE: num, name	see FILES	see FILES
FILE NAME MAX	installation determined	--------	8 char	6 characters not including extensions		6 char	6 char
FILES	--------	--------	see FILE	see OPEN	Files name, ...	FILES name	FILES name
FNEND	--------	--------	FNEND	--------	FNEND	--------	--------
FOR **	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3
GETPTR	--------	--------	--------	--------	LCW(exp)	--------	--------

**Commands and elements that can be used.

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
ENTER	--------	--------	--------	--------	--------	--------	--------
EQUAL	=	=	=	=	EQU(exp1, exp2)	=	=
EQUIVALENCE	--------	--------	--------	EQV	EQU(exp1, exp2)	--------	--------
EXCHANGE	--------	--------	--------	FILE exp	EXCHANGE v1, v2 or v1 : = v2	--------	OPEN name to: str, PRINT, ON
EXCLUSIVE OR	--------	--------	--------	EOR	XOR(exp1, exp2)	--------	--------
EXPONENTIATION **	↑ or **	↑	↑ or **	↑	↑ or **	↑	↑ or **
FIELD	see IMAGE	see PRINT USING	see PRINT USING	see PRINT USING	--------	FIELD (spec1, spec2...spec n	see PRINT USING
FILE	--------	--------	FILE #N, str... FILE: N, str...	see OPEN	see OPEN	see ASSIGN	see OPEN
FILE NAME MAX	3 char	6 char		6 char	12 char	7 char	11 char
FILES	--------	FILES name, ...	FILES name, ...	see OPEN	see OPEN	see ASSIGN	see OPEN
FNEND	--------	--------	FNEND	FNEND	FNEND	--------	--------
FOR **	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3 FOR var = exp1 TO exp3 BY exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3	FOR var = exp1 TO exp2 STEP exp3
GETPTR	--------	--------	--------	--------	--------	GETPTR num, var1, var2	--------

3B

**Commands and elements that can be used.

3C

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
ENTER	------	------	------	------	------	------	ENTER # var1, ... ENTER var1,...
EQUAL	=	=	=	=	=	=	=
EQUIVALENCE	------	------	------	------	------	------	------
EXCHANGE	------	------	------	------	------	------	------
EXCLUSIVE OR	------	------	------	------	------	------	------
EXPONENTIATION**	↑ or **	↑	↑ or ^	or **	↑	↑ or **	↑
FIELD	see PRINT USING	------	FMT or see PRINT USING	------	------	------	see IMAGE
FILE	FILE name,num not collect	------	see FILES	------	------	see FILES	see FILES
FILE NAME MAX	8 characters	------	letter or letter num	7 characters	------	17 characters	6 characters
FILES	see FILE	------	FILES name1,...	------	------	FILES name1,...	FILES name,...
FNEND	------	------	------	------	------	------	------
FOR**	FOR var=exp1 TO exp2 STEP exp3	FOR var=exp1 TO exp2 STEP exp3	FOR var=exp1 TO exp2 STEP exp3	FOR var=exp1 TO exp2 STEP exp3	FOR var=exp1 TO exp2 STEP exp3	FOR var=exp1 TO exp 2 STEP exp3	FOR var=exp1 TO exp 2 STEP exp3
GETPTR	------	------	------	------	------	------	------

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC	3D
ENTER	--------	--------	--------	--------	--------	--------	--------	
EQUAL**	=	= or \EQ or ←	=	=	=	=	=	
EQUIVALENCE	--------	--------	--------	--------	--------	--------	--------	
EXCHANGE	--------	--------	--------	--------	--------	--------	--------	
EXCLUSIVE OR	--------	--------	--------	--------	--------	--------	--------	
EXPONENTIATION**	↑	**	**	**	↑ integer < = 9	↑	↑	
FIELD	--------	--------	--------	--------	see IMAGE	--------	--------	
FILE	--------	see FILES	--------	--------	see OPEN	--------	--------	
FILE NAME MAX	--------	6 characters	--------	--------	6 characters	--------	6 characters program files only	
FILES	--------	--------	--------	--------	see OPEN	--------	--------	
FNEND	--------	--------	--------	--------	--------	--------	--------	
FOR**	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	
GETPTR	--------	--------	--------	--------	KEY	--------	--------	

**Commands and elements that can be used.

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16
ENTER	--------	--------	--------	--------	--------	--------	--------
EQUAL**	=	=	=	=	=	=	=
EQUIVALENCE	--------	--------	--------	EQU	--------	--------	--------
EXCHANGE	--------	--------	--------	--------	--------	--------	--------
EXCLUSIVE OR	--------	--------	--------	XOR	--------	--------	XOR
EXPONENTIATION**	↑ or **	↑	↑	↑ or **	↑	↑	↑ or !
FIELD	see IMAGE	--------	--------	see PRINT USING	--------	--------	--------
FILE	ALLOCATE	see FILES	see OPEN	see OPEN	--------	--------	--------
FILE NAME MAX	8 characters	6 characters	9 characters	9 characters	--------	--------	--------
FILES	see FILE	FILES name 1; name 2; ...	see OPEN	see OPEN	--------	--------	--------
FNEND	--------	--------	--------	see RETURN	--------	--------	--------
FOR**	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3 max 26	FOR var=exp 1 TO exp 2 STEP num	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3	FOR var=exp 1 TO exp 2 STEP exp 3
GETPTR	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
ENTER	--------	--------	--------	--------	--------	--------	--------
EQUAL**	EQU (exp 1, exp 2) or =	=	=	=	EQ or =	=	=
EQUIVALENCE	EQV (exp 1, exp 2)	--------	--------	--------	--------	--------	--------
EXCHANGE	EXCHANGE var 1, var 2 or var 1 == var 2	--------	--------	--------	--------	--------	--------
EXCLUSIVE OR	XOR (exp 1, exp 2)	--------	--------	--------	--------	--------	--------
EXPONENTIATION**	↑ or Δ or **	↑ or **	↑	↑	↑ or **	↑	↑ or **
FIELD	see PRINT USING	FMT specifications	--------	see IMAGE	see IMAGE	--------	--------
FILE	see OPEN	see FILES	--------	FILE # num, name	FILE # num, name, password FILE: num, name, password	see FILES	--------
FILE NAME MAX	12 characters	6 characters	--------	6 characters	12 characters	6 characters	--------
FILES	see OPEN	FILES name 1, ... name n	--------	FILES name 1; ... name n	FILES name; password; ... or FILES options	FILES name, ...	--------
FNEND	FNEND	--------	--------	--------	FNEND	--------	FNEND
FOR**	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3 or FOR var = exp 1, exp 2, exp 3	FOR var = exp 1 TO exp 2 STEP exp 3 or FOR var = exp 1, exp 2, exp 3	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3 FOR var = exp 1 TO exp 2 BY exp 3
GETPTR	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

431

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200		3G
ENTER	--------	--------	ENTER # var ENTER num 1, num 2, num 3 ENTER * num 1, num 2, num 3, var	--------	--------	--------		
EQUAL**	=	=	=	=	=	=		
EQUIVALENCE	--------	--------	--------	--------	--------	--------		
EXCHANGE	--------	--------	--------	--------	--------	--------		
EXCLUSIVE OR	--------	--------	--------	--------	--------	--------		
EXPONENTIATION**	↑	↑	^ or ↑	↑	↑	↑		
FIELD	--------	--------	see PRINT USING	see IMAGE	see IMAGE	see IMAGE		
FILE	--------	--------	see FILES	see FILES	see FILES	see ASSIGN		
FILE NAME MAX	--------	--------	depends on installation	8 characters	6 characters	8 characters		
FILES	--------	--------	FILES * num FILES name	FILES options	FILES name 1; name 2; ...	see ASSIGN		
FNEND	--------	--------	FNEND	--------	--------	--------		
FOR**	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3	FOR var = exp 1 TO exp 2 STEP exp 3		
GETPTR	--------	--------	REC	--------	LOC (exp)	--------		

**Commands and elements that can be used.

432

4A

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C
GOSUB**	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #
GOSUB OF (computed)	--------	--------	see ON	see ON	--------	GOSUB exp OF n1, n2, ...	GOSUB exp OF n1, n2, ...
GOTO**	GOTO line #	GOTO num GO TO num	GO TO line #	GO TO line #	GO TO num	GOTO line # GO TO line #	GO TO line # GOTO exp line #
GOTO OF (computed)	--------	--------	--------	see ON	--------	GOTO exp OF n1, n2, ...	GOTO exp OF n1, n2, ...
GREATER**	>	>	>	>	> or .GT.	>	>
GREATER EQUAL**	> = or = >	> = or = >	> = or = >	> =	> = or .GE.	> =	> =
HOLD	--------	--------	--------	--------	--------	--------	--------
IF**	IF exp1 op exp2 THEN line #	IF exp1 op exp2 THEN line #	IF exp1 op exp2 THEN line # or IF exp1 op exp2 GOTO line #	IF exp1 op exp2 THEN line # IF exp1 op exp2 GOTO line # IF exp1 op exp2 GOSUB line # IF exp1 op exp2 THEN statement	IF exp1 op exp2 THEN line # IF exp1 op exp2 GOTO line #	IF exp1 op exp2 THEN line #	IF exp1 op exp2 THEN line #
IF END #	--------	--------	IF END # exp THEN line #	IF EOF (num) THEN line #	IF END # exp THEN line # IF END # exp: THEN line #	IF END # exp THEN line #	IF END # exp THEN line #
IF MORE #	--------	--------	IF MORE # exp THEN line #	--------	IF MORE # exp THEN line # IF MORE # exp: THEN line #	--------	--------

**Commands and elements that can be used.

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
GOSUB**	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #
GOSUB OF (computed)	--------	GOSUB exp OF line #, ...	--------	--------	--------	--------	--------
GOTO**	GOTO line #	GOTO line #	GO TO line #	GOTO line #	GO TO line # GO TO * num	GO TO line # GOTO exp, line #	GOTO line #
GOTO OF (computed)	--------	GOTO exp of line #, ...	--------	--------	--------	--------	--------
GREATER**	>	>	>	>	GTR(exp1, exp2)	>	>
GREATER EQUAL**	> =	> =	> =	> =	GEQ(exp1, exp2)	> =	> = or = >
HOLD	--------	--------	--------	--------	--------	HOLD μ	--------
IF**	IF exp1 op exp2 THEN line # IF exp1 op exp2 GOTO line #	IF exp1 op exp2 THEN line #	IF exp1 op exp2 THEN line # IF exp1 op exp2 THEN line #	IF exp1 op exp2 THEN statement IF exp1 op exp2 THEN line # IF exp1 op exp2 GOTO line #	IF exp1 op exp2 THEN line # IF exp1 op exp2 THEN statement IF exp1 op exp2 GOTO line # IF exp1 op exp2 GOSUB line #	IF exp1 op exp2 THEN line # or IF exp1 op exp2 GOTO line #	IF exp1 op exp2 THEN line #
IF END #	--------	IF END # exp THEN line #	IF END #var THEN line # IF END: var THEN line #	--------	see ON	--------	--------
IF MORE #	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
GOSUB**	GOSUB line#	GOSUB line#	GOSUB line#	GOSUB line#	GOSUB line#	GOSUB line# GO SUB line#	GOSUB line #
GOSUB OF (computed)	------	------	GOSUB(line#, ... line#)exp	------	------	------	GOSUB exp OF line#,...
GOTO**	GOTO line#	GOTO line#	GO TO line#	GO TO line#	GO TO line# GOTO line#	GO TO line# GOTO line#	GO TO line# GOTO line#
GOTO OF (computed)	GOTO line#,... line# ON exp	------	GO TO(line#, ...line#) exp or see ON	------	------	------	GOTO exp OF line#,...
GREATER**	>	>	>	>	>	>	>
GREATER EQUAL**	>= or =>	>=	=>	>= or =>	>= or =>	>= or =>	>=
HOLD	------	------	------	------	------	------	------
IF**	IF exp1 op exp2 THEN line# IF exp1 op exp2 GOTO line #	IF exp1 op exp2 THEN line# IF exp1 op exp2 GOTO line #	IF exp1 op exp2 THEN line# IF exp1 op exp2 GO TO line #	IF exp1 op exp2 THEN line #	IF exp1 op exp2 THEN line# IF exp1 op exp2 GO TO line #	IF exp1 op exp2 THEN line#	IF exp1 op exp2 THEN line#
IF END #	------	------	IF END # name GO TO line# IF END # name THEN line#	see NODATA	------	IF END # exp THEN line#	IF END # exp THEN line#
IF MORE #	------	------	------	------	------	------	------

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC
GOSUB**	GOSUB line #	GOSUB line #	GOSUB line#	GOSUB line#	GOSUB line#	GOSUB line#	GOSUB line#
GOSUB OF (computed)	--------	--------	--------	--------	--------	--------	GOSUB exp OF line#,...GOSUB line#, param, ...
GOTO**	GO TO line #	GO TO line #	GO TO line #	GO TO line #	GO TO line #	GO TO line #	GO TO line #
GOTO OF (computed)	--------	--------	--------	--------	--------	--------	GO TO exp OF line#,...
GREATER**	>	\GT or >	>	>	>	>	>
GREATER EQUAL**	> = or = >	\GE or >= or => or ≥	>=	>=	> = or = >	>=	>=
HOLD	--------	--------	--------	--------	--------	--------	--------
IF**	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp statement	IF exp op exp THEN line #	IF exp op exp THEN line #
IF END #	--------	--------	--------	--------	see READ # or WRITE #	--------	--------
IF MORE #	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16
GOSUB**	GOSUB line #	GOSUB line # no recursion	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #
GOSUB OF (computed)	--------	--------	ON exp GOSUB line #, ...	ON exp GOSUB line #, ...	--------	--------	--------
GOTO**	GO TO line #	GO TO line #	GOTO line #	GOTO line #	GOTO line #	GOTO line #	GOTO line #
GOTO OF (computed)	GO TO line #, ... line # ON exp	ON exp GO TO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...
GREATER**	>	>	>	>	>	>	>
GREATER EQUAL**	>=	>=	>=	>=	>= or =>	>= or =>	>=
HOLD	--------	--------	--------		--------	--------	--------
IF**	IF exp op exp THEN line # IF exp op exp GO TO line #	IF exp op exp THEN line #	IF exp op exp THEN line # IF exp op exp GOTO line # IF exp op exp GOSUB line #	IF exp op exp THEN line # IF exp op exp GOTO line # IF exp op exp GOSUB line # IF exo op exp THEN statement	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp THEN line # IF exp op exp GOTO line #
IF END #	--------	IF END # exp THEN line num	--------	see ENDFILE	--------	--------	--------
IF MORE #	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

4E

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
GOSUB**	GOSUB line #	GO SUB line #	GO SUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #
GOSUB OF (computed)	--------	--------	--------	--------	-------	GOSUB exp OF line #, ...	--------
GOTO**	GOTO line # or GOTO * num	GOTO line #	GOTO line #	GOTO line #	GOTO line #	GOTO line #	GOTO line # GOTO * ± num
GOTO OF (computed)	ON exp THEN line #, ... ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ... ON exp THEN line #, ...	GOTO exp OF line #, ...	ON exp THEN line #, ... ON exp GOTO line #, ...
GREATER**	GTR (exp 1, exp 2) or >	>	>	>	> or GT	>	>
GREATER EQUAL**	GEQ(exp 1, exp 2)or > = or = >	> = or = >	> = or = >	> = or = >	GE or > = or = >	> =	> = or = >
HOLD	--------	--------	--------	--------	--------	--------	--------
IF**	IF exp op exp THEN line # IF exp op exp GOTO line # IF exp op exp THEN statement	IF exp op exp THEN line # IF exp op exp GOTO line # IF exp op exp line # 1, line # 2, line # 3	IF exp op exp THEN line # IF exp op exp GOTO line # IF exp op exp THEN statement IF exp op exp line # 1, line # 2, line # 3	IF exp op exp THEN line #	IF exp op exp THEN line # IF exp op exp GOTO line #	IF exp op exp THEN line #	IF exp op exp THEN line # IF exp op exp GOTO line #
IF END #	see ENDFILE	IF END # exp THEN line #	--------	IF END # num, THEN line #	IF END # num THEN line # IF END # num GOTO line #	IF END # exp THEN line #	--------
IF MORE #	--------	--------	--------	IF MORE # num, THEN line #	IF MORE # num, THEN line # IF MORE # num, GOTO line #	--------	--------

**Commands and elements that can be used.

4F

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200	4G
GOSUB**	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	GOSUB line #	
GOSUB OF (computed)	--------	--------	GOSUB exp OF line #, ...	--------	--------	--------	
GOTO**	GOTO line #	GOTO line #	GOTO line #	GOTO line #	GO TO line #	GOTO line #	
GOTO OF (computed)	ON exp GO TO line #, ...	--------	GOTO exp OF line #, ...	GOTO line #, ... line # ON exp	ON exp GOTO line #, ...	--------	
GREATER**	>	>	>	>	>	>	
GREATER EQUAL**	>=	>= or =>	>=	>=	>=	>=	
HOLD	--------	--------	--------	--------	--------	--------	
IF**	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp THEN line #	IF exp op exp THEN line #	
IF END #	--------	--------	IF END # exp THEN line # ON END # exp THEN line #	IF END # num THEN line #	IF END # exp THEN line # IF END: exp THEN line #	--------	
IF MORE #	--------	--------	--------	--------	--------	--------	

**Commands and elements that can be used.

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C	SA
IF-THEN-ELSE	--------	--------	--------	--------	--------	--------	--------	
IMAGE	--------	IMAGE	see USING	--------	--------	see USING	see USING	
IMPLICATION	--------	--------	--------	--------	--------	-------	-------	
INPUT**	INPUT var1, ...	INPUT ARRAY INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	
INPUT FROM	--------	--------	--------	--------	--------	--------	--------	
INPUT LINE	--------	--------	see INPUT	-------	--------	--------	--------	
INPUT (-) list	--------	--------	--------	-------	--------	--------	--------	
INPUT #	--------	--------	INPUT # exp: var1, var2, ...	INPUT FILE (e.p) list	INPUT # exp, list INPUT # exp: list	--------	--------	
INPUT " ", var, ...	--------	--------	--------	--------	--------	--------	--------	
INTEGER NAME	--------	--------	--------	--------	--------	--------	-- -----	
KILL	--------	--------	see SCRATCH	--------	--------	not collect	not collect	
LARGEST # **	E337	7.2 E75	1.70141 E38	7.2 E75	1.70141 E38	E38	E38	
LESS**	<	<	<	<	< or .LT.	<	<	
LESS EQUAL**	< = or = <	< = or = <	< = or = <	< =	< = or .LE.	< =	< =	
LET **	LET var1 = var2 = ... = exp or no LET	LET var1, ... var n = exp or no LET	LET var1 = var2 = ... exp	LET var = exp or no LET	LET var1 = var2 = var3 ... = exp	LET var1 = var2 = ... = exp or no LET	LET var1 = var2 = ... exp or no LET	

**Commands and elements that can be used.

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
IF-THEN-ELSE	--------	--------	--------	IF statement op statement THEN statement ELSE statement	--------	--------	--------
IMAGE	IMAGE	see PRINT USING	see PRINT USING	see PRINT USING	--------	see FIELD	see PRINT USING
IMPLICATION	--------	--------	--------	IMP	IMP(exp1, exp2)	--------	--------
INPUT**	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...	INPUT var1, ...
INPUT FROM	--------	--------	--------	--------	INPUT FROM num: list	--------	--------
INPUT LINE	--------	--------	--------	INPUT LINE string variable	--------	--------	--------
INPUT ( ) list	--------	--------	--------	--------	see INPUT FROM	INPUT (61, n) var1, ...	--------
INPUT #	GET name, list	--------	INPUT #N, list INPUT: N, list	INPUT # exp, var1, var2 ...	see INPUT FROM	--------	INPUT: str; key, list
INPUT " ", var, ...	--------	--------	--------	INPUT "str"; var ...	--------	--------	--------
INTEGER NAME	--------	--------	--------	numeric name followed by %	--------	--------	--------
KILL	--------	KILL name	--------	KILL string	--------	see SCRATCH	--------
LARGEST #***	7.2 E75	E38	1.7 E38	1 [illegible] E38	E38	E99	7.237 E75
LESS**	<	<	<	<	LSS(exp1, exp2)	<	<
LESS EQUAL**	< =	< =	< =	< =	LEQ(exp1, exp2)	< =	< = or = <
LET**	LET var1, ... = exp	LET var1 = ... var n = exp	LET var1 = ... = exp or no LET	LET var1, var2, ... var n = exp or no LET	LET var1 = var2 = ... exp	LET var1, var2, ... var n = exp or no LET	LET var1, ... = exp or no LET

**Commands and elements that can be used.

58

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F	SC
IF-THEN-ELSE	------	------	------	------	------	------	------	
IMAGE	see PRINT USING	------	FMT	------	------	------	IMAGE	
IMPLICATION	------	------	------	------	------	------	------	
INPUT**	INPUT var1,...	INPUT var1,...	INPUT var1,...	INPUT var1,...	INPUT var1,...	INPUT var1,...	INPUT var1,...	
INPUT FROM	------	------	------	------	------	------	------	
INPUT LINE	------	------	------	------	------	------	------	
INPUT ( ) list	see READ #	------	see READ #	see INPUT #	------	------	see READ #	
INPUT #	see READ #	------	see READ #	INPUT FILE (name) var1,...	------	------	see READ #	
INPUT " ", var, ...	see READ #	------	see READ #	see INPUT #	------	------	see READ #	
INTEGER NAME	------	------	------	------	------	------	------	
KILL	------	------	------	------	------	------	KILL-name KIL-name	
LARGEST #**	1E 75	1E 615	1E 616	1E 337	7.2 E 75	1E 47	1E 38	
LESS**	<	<	<	<	<	<	<	
LESS EQUAL**	< = or = <	< =	< =	< = or = <	< = or = <	< = or = <	< =	
LET**	LET var1,... var n = exp or no LET	LET var = exp	LET var1=var2= ...var n= exp or no LET	LET var1=var2= ...var n=exp or no LET	LET var1=exp	LET var1=... =var n=exp	LET var1=... var n=exp	

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

	NCR CENTURY 100 BASIC I	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC
IF-THEN-ELSE	--------	--------	--------	--------	--------	--------	--------
IMAGE	--------	--------	--------	--------	TABLE	--------	--------
IMPLICATION	--------	--------	--------	--------	--------	--------	--------
INPUT**	INPUT var1,...	INPUT var1,...	INPUT var1,...	INPUT var1,...	INPUT exp1,...	INPUT var1,...	INPUT var1,...
INPUT FROM	--------	--------	--------	--------	--------	--------	--------
INPUT LINE	--------	--------	--------	--------	--------	--------	--------
INPUT ( ) list	--------	see INPUT #	--------	--------	see INPUT #	--------	--------
INPUT #	--------	INPUT FILE name var list INPUT#exp, var list	--------	--------	INPUT(num, IND =exp, ERR=num, END=num)exp1, exp2, ...	--------	--------
INPUT " ", var, ...	--------	see INPUT #	--------	--------	see INPUT #	--------	--------
INTEGER NAME	--------	--------	--------	--------	--------	--------	--------
KILL	--------	--------	--------	--------	ERASE name	--------	--------
LARGEST # **	1E 99	4.314E 68	1E99	1E99	1E99	1.67E 73	1E99
LESS **	<	\LT or <	<	<	<	<	<
LESS EQUAL **	< = or = <	\LE or < = or = < or ≤	<=	<=	< = or = <	<=	<=
LET **	LET var=exp	LET var1=var2= ...var n= exp or no LET	LET var=exp or no LET	Let var=exp or no LET	LET var=exp	LET var=exp	LET var1=var2= ...var n = exp

**Commands and elements that can be used.

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16	SE
IF-THEN-ELSE	--------	--------	--------	IF exp op exp THEN statement ELSE statement	--------	--------	--------	
IMAGE	str or image	--------	--------	see PRINT USING	--------	--------	--------	
IMPLICATION	--------	--------	--------	IMP	--------	--------	--------	
INPUT**	INPUT var 1, ...	INPUT var 1, ...	INPUT var 1, ...	INPUT var 1, ... or DISPLAY or ACCEPT	INPUT var 1, ...	INPUT var 1, ...	INPUT var 1, ... or INPUT $ input device $ var 1, ...	
INPUT FROM	--------	--------	--------	INPUT FROM exp: var 1, ...	--------	--------	--------	
INPUT LINE	--------	--------	--------	--------	--------	--------	--------	
INPUT ( ) list	--------	--------	--------	see INPUT FROM	--------	--------	--------	
INPUT #	--------	--------	INPUT FILE var 1, ...	see INPUT FROM	--------	--------	--------	
INPUT " ", var, ...	--------	--------	see INPUT #	see INPUT FROM	--------	--------	--------	
INTEGER NAME	--------	--------	--------	num name	--------	--------	--------	
KILL	--------	--------	--------	--------	--------	--------	--------	
LARGEST #**	1E99	5.78960E76	5E76	5E76	9.23E18	9.23E18	9.23E18	
LESS**	<	<	<	<	<	<	<	
LESS EQUAL**	< =	< =	< =	< =	< = or = <	< = or = <	< =	
LET**	LET var 1, ... var n = exp or no LET	LET var = exp	LET var 1=var 2= ... var n = exp or no LET	LET var 1, ... =exp or LET var 1 + var 2 ... + exp or no LET	LET var 1 = var 2 = ... var n = exp or no LET	LET var 1 = var 2 = ... var n = exp or no LET	LET var 1 = var 2 = ... var n = exp	

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
IF-THEN-ELSE	IF exp op exp THEN statement ELSE statement	--------	--------	--------	--------	--------	--------
IMAGE	see PRINT USING	see FIELD	--------	: format	: format	--------	--------
IMPLICATION	IMP(exp 1, exp 2)	--------	--------	--------	--------	--------	--------
INPUT**	INPUT var 1, ... var n	INPUT var 1, ..., var n	INPUT var 1, ..., var n	INPUT var 1, ..., var n	INPUT var 1, ..., var n	INPUT var 1, ..., var n	INPUT var 1, ..., var n
INPUT FROM	INPUT FROM exp: var 1, ... var n	--------	--------	--------	--------	--------	--------
INPUT LINE	--------	--------	--------	--------	--------	--------	--------
INPUT ( ) list	see INPUT FROM	--------	--------	see INPUT #	see INPUT #	--------	--------
INPUT #	see INPUT FROM	--------	--------	INPUT: name: var 1, ... INPUT # num, var 1, ...	INPUT # num, list	--------	--------
INPUT " ", var, ...	see INPUT FROM	--------	--------	see INPUT #	see INPUT #	see INPUT #	--------
INTEGER NAME	--------	--------	--------	--------	--------	--------	--------
KILL	--------	--------	--------	--------	--------	KILL name	--------
LARGEST #***	1E38	1E38	1E38	5.7896E76	1E38	1E38	1E38
LESS**	LSS(exp 1, exp 2) or <	<	<	<	LT or <	<	<
LESS EQUAL**	LEQ(exp 1, exp 2) or < = or = <	< = or = <	< = or = <	< = or = <	LE or < = or = <	< =	< = or = <
LET**	LET var 1 = ... var n = exp or no LET	LET var = exp or no LET	LET var 1, var 2, ... var n = exp or no LET	LET var = exp or no LET	LET var 1 = var 2 = ... var n = exp or no LET	LET var 1 = var 2 = ... var n = exp or no LET	LET var 1 = var 2 = ... var n = exp or no LET

**Commands and elements that can be used.

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200		5G
IF-THEN-ELSE	--------	--------	IF exp op exp THEN statement ELSE statement	--------	--------	--------		
IMAGE	--------	--------	IMAGE formats	% formats	: formats	% formats		
IMPLICATION	--------	--------	--------	--------	--------	--------		
INPUT**	INPUT var 1, ... var n	INPUT var 1, ... var n	INPUT var 1, ... var n	INPUT var 1, ... var n	INPUT var 1, ... var n	INPUT var 1, ... var n		
INPUT FROM	--------	--------	--------	--------	--------	--------		
INPUT LINE	--------	--------	--------	--------	--------	--------		
INPUT ( ) list	--------	--------	--------	--------	--------	--------		
INPUT #	--------	--------	--------	--------	--------	--------		
INPUT " ", var, ...	--------	--------	--------	--------	--------	--------		
INTEGER NAME	--------	--------	letter	--------	--------	--------		
KILL	--------	--------	--------	--------	--------	--------		
LARGEST #**	1E37	1E99	1E77	1E63	5.78960E76	1E100		
LESS**	<	<	<	<	<	<		
LESS EQUAL**	< =	< = or = <	< =	< =	< =	< =		
LET**	LET var = exp	LET var = exp	LET var 1 = var 2 = ... var n = exp or no LET	LET var 1, var 2, ... var n = exp or no LET	LET var = exp	LET var 1, var 2, ... var n = exp or no LET		

**Commands and elements that can be used.

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C
LINE #**	0 to 99999	1 to 999	1 to 99999	1 to 9999	1 to 99999	1 to 9999	1 to 9999
LINPUT	--------	--------	LINPUT # N: list of str var	--------	--------	--------	--------
LONGEST STRING **	72	15	4095	256	119	72	72
MARGIN	--------	--------	MARGIN # N: exp	--------	MARGIN # exp, exp MARGIN # N, MARGIN # exp: exp	--------	--------
MAT **	MAT	MAT	MAT	MAT	MAT	MAT	MAT
MAXIMUM	--------	--------	--------	--------	--------	MAX	MAX
MAX ARRAY SIZE		500 elements		1024 elements		2500 elements	2500 elements
MAX NESTING LOOP	10		6	4	20	9	9
MAX # OF DIM IN AN ARRAY	3	2	2	2	2	2	2
MINIMUM	--------	--------	--------	--------	--------	MIN	MIN
MULTIPLE STATEMENTS	--------	--------	--------	--------	--------	--------	--------
NAME AS	--------	--------	--------	--------	--------	--------	--------
NEXT**	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var
NODATA	NODATA num NODATA FILE (name)	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2,0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
LINE #**	1 to 99999	1 to 9999	1 to 99999	1 to 32767	1 to 99999	1 to 99999	1 to 99999
LINPUT	--------	--------	--------	--------	--------	--------	--------
LONGEST STRING **	18	198 char	size of core	size of core	512	80	22
MARGIN	--------	--------	MARGIN exp	--------	--------	--------	--------
MAT **	MAT	MAT	MAT	MAT	MAT	MAT	MAT
MAXIMUM	--------	MAX	--------	--------	--------	--------	--------
MAX ARRAY SIZE							
MAX NESTING LOOP	15			Depends on storage	32		26
MAX # OF DIM IN AN ARRAY	2	2	2	2	2	2	2
MINIMUM	--------	MIN	--------	--------	--------	--------	--------
MULTIPLE STATEMENTS	--------	--------	--------	separated by :	--------	--------	--------
NAME AS	--------	--------	--------	NAME str AS str protection	--------	--------	--------
NEXT **	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT	NEXT var
NODATA	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
LINE #**	1 to 99999	1 to 2046	1 to 99999	0 to 99999	1 to 9999	0 to 99999	1 to 9999
LINPUT	------	------	------	------	------	------	------
LONGEST STRING **	18 char	------	63 char	72 char	14 char	15 char	72 char
MARGIN	------	------	------	------	------	------	------
MAT **	MAT	------	MAT	MAT	------	------	MAT
MAXIMUM	------	------	MAX(var1,...)	------	------	------	MAX
MAX ARRAY SIZE	28,668 bytes	core restricted	core restricted	core restricted	4096 by 4096	4095 by 4095	4900 elements
MAX NESTING LOOP	15	8	10	10	10	no limit	9
MAX # OF DIM IN AN ARRAY	2	2	3	3	2	2	2
MINIMUM	------	------	MIN(var1,...)	------	------	------	MIN
MULTIPLE STATEMENTS	------	------	separated by \	------	------	------	------
NAME AS	------	------	------	------	------	------	------
NEXT **	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var
NODATA	------	------	------	NODATA name NODATA FILE (name) num	------	------	see IF END

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUCHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC
LINE # **	999999	99999999	9999	9999	1 to 9999	1 to 9999	1 to 9999999
LINPUT	--------	--------	--------	--------	--------	--------	--------
LONGEST STRING **	--------	15 characters	15 characters (in PRINT)	15 characters (in PRINT)	core determined	--------	--------
MARGIN	--------	--------	--------	--------	--------	--------	--------
MAT **	--------	MAT	MAT	MAT	--------	--------	MAT
MAXIMUM	--------	--------	--------	--------	--------	--------	--------
MAX ARRAY SIZE	1512 elements	1023 by 1024	1000 elements	1000 elements	999 elements	?	255 by 255
MAX NESTING LOOP	10	core dependent	5	5	5	?	?
MAX # OF DIM IN AN ARRAY	2	2	2	2	3	2	3
MINIMUM	--------	--------	--------	--------	--------	--------	--------
MULTIPLE STATEMENTS	--------	--------	--------	--------	--------	--------	--------
NAME AS	--------	--------	--------	--------	--------	--------	--------
NEXT **	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var
NODATA	--------	--------	----- --	--------	--------	--------	--------

**Commands and elements that can be used.

60

450

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16
LINE #**	0 to 9999	1 to 99999	0 to 99999	1 to 99999	1 to 9999	1 to 9999	1 to 9999
LINPUT	--------	--------	--------	--------	--------	--------	--------
LONGEST STRING**	18 characters	15 characters	core dependent	core dependent	--------	72 characters	72 characters
MARGIN	--------	--------	--------	--------	--------	--------	--------
MAT**	MAT	MAT	MAT	MAT	--------	MAT	MAT
MAXIMUM	--------	--------	--------	MAX(var 1, ...) is function	--------	--------	--------
MAX ARRAY SIZE	set by system definition	2074	core dependent	core dependent	32,767	32,767	core dependent
MAX NESTING LOOP	9	26	core dependent	core dependent	no limit	no limit	core dependent
MAX # OF DIM IN AN ARRAY	2	2	2	limited by statement length	limited by line length	limited by line length	limited by line length
MINIMUM	--- ----	--------	--------	MIN (var 1, ...) is function	--------	--------	--------
MULTIPLE STATEMENTS	--------	--------	--------	YES \ or, or;	--------	--------	--------
NAME AS	--------	--------	--------	--------	--------	--------	--------
NEXT**	NEXT var	NEXT var	NEXT var	NEXT var or NEXT line # or NEXT var, ...	NEXT var	NEXT var	NEXT var
NODATA	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
LINE #**	0 to 99999	1 to 32767	1 to 9999	1 to 99999	1 to 99999999	1 to 9999	0 to 99999
LINPUT	--------	--------	--------	--------	--------	--------	--------
LONGEST STRING**	511 characters	depends on core	--------	132 characters	132 characters	72 characters	60 characters
MARGIN	-------	-------	-------	MARGIN # num, exp	MARGIN # num, exp	--------	--------
MAT**	MAT	MAT	--------	MAT	MAT	MAT	MAT
MAXIMUM	MAX(exp 1, exp 2) function	--------	--------	--------	--------	MAX	--------
MAX ARRAY SIZE	determined by installation	depends on available core	depends on core	2000 elements in core 22000 on disc	depends on core available	2000 elements	depends on core
MAX NESTING LOOP	32	8	depends on installation	6	26	6	32
MAX # OF DIM IN AN ARRAY	4	2	depends on line length	depends on line length	2	2	2
MINIMUM	MIN (exp 1, exp 2) function	--------	--------	--------	--------	MIN	--------
MULTIPLE STATEMENTS	--------	--------	separated by:	--------	separated by \	--------	--------
NAME AS	--------	--------	--------	--------	--------	--------	--------
NEXT **	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var
NODATA	--------	--------	--------	--------	--------	--------	--------

**Commands and elements that can be used.

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200	
LINE #**	1 to 9999	1 to 999998	1 to 9999	1 to 9999	1 to 99999	1 to 9999	
LINPUT	--------	--------	LINPUT string variable	--------	--------	--------	
LONGEST STRING**	--------	--------	255 characters	18 characters	15 characters	64 characters	
MARGIN	--------	--------	--------	--------	--------	--------	
MAT**	--------	--------	MAT	MAT	MAT	--------	
MAXIMUM	--------	--------	MAX	--------	--------	--------	
MAX ARRAY SIZE	depends on core	1512 elements	depends on core	dimensions ≤ 255	2074 elements	dimensions ≤ 255	
MAX NESTING LOOP	depends on core	10	depends on core	depends on core	26	depends on core	
MAX # OF DIM IN AN ARRAY	2	2	2	2	2	2	
MINIMUM	--------	--------	MIN	--------	--------	--------	
MULTIPLE STATEMENTS	--------	--------	in LET statement separated by ,	separated by :	--------	separated by :	
NAME AS	--------	--------	--------	--------	--------	--------	
NEXT**	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	NEXT var	
NODATA	--------	--------	--------	--------	--------	--------	

**Commands and elements that can be used.

66

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C
NOT	--------	--------	--------	--------	--------	NOT	NOT
NOT EQUAL **	<> or ><	¬= or <>	<>	<>	<> or .NE.	<> or #	<> or #
NUMERIC VARIABLE NAME **	letter or letter digit	letter or letter alphameric	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit
ON	ON exp GOTO num, ...	--------	ON exp GOTO line #, line # ... ON exp GOSUB line #, line # ...	ON exp GOTO line #, ... ON exp THEN line #, ... ON exp GOSUB line #, ...	ON exp GOTO line #, ...	--------	--------
ON ERROR GOTO	--------	--------	--------	--------	--------	--------	--------
OPEN	see READ or WRITE FILE	--------	see FILE	OPEN FILE num, num, name	--------	OPEN-name, num OPE -name, num	OPEN-name, num OPE -name, num, num
OR	--------	--------	--------	--------	--------	OR	OR
PAUSE	--------	--------	--------	--------	--------	--------	--------
PLOT	--------	--------	--------	--------	--------	--------	--------
PRINT**	PRINT list of exp	PRINT list	PRINT list of exp	PRINT list of exp or; list of exp	PRINT list	PRINT list of exp	PRINT list of exp
PRINT #	see WRITE FILE	--------	PRINT # exp: list of exp	PRINT FILE [exp] list of exp	PRINT # exp, list PRINT # exp: list	PRINT # exp: list, END PRINT # expX: list, END	PRINT # expl; PRINT # exp: list, END
PRINT # USING	--------	--------	PRINT # exp: USING str exp. list	PRINT FILE (exp) USING "string", list	PRINT # exp, USING str, list	--------	--------

**Commands and elements that can be used.

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
NOT	--------	NOT	--------	NOT	NOT(exp)	--------	--------
NOT EQUAL **	<>	<>	<>	<>	NEQ(exp1, exp2)	<>	<> or ><
NUMERIC VARIABLE NAME**	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit
ON	GOTO varl, ... var n ON exp	--------	ON exp GO TO line #, ... ON exp THEN line #, ...	ON exp GOTO line #, line # ... ON exp GOSUB line #, line # ...	ON ENDFILE num GOTO line #	ON exp GOTO line #, line # ...	ON exp GOTO line #, ...
ON ERROR GOTO	--------	--------	--------	ON ERROR GO TO	--------	--------	--------
OPEN	--------	OPEN-name, num, num	--------	OPEN name FOR INPUT OUTPUT as	OPEN name SYMBOLIC INPUT BINARY OUTPUT,num	OPEN u, name, m	OPEN name, TO: str, GET, T FILE
OR	--------	OR	--------	OR	IOR(exp1, exp2)	--------	--------
PAUSE	PAUSE	--------	--------	--------	--------	--------	PAUSE
PLOT	--------	--------	--------	--------	--------	PLOT exp1 = "exp2"	--------
PRINT**	PRINT list of exp	PRINT list	PRINT list	PRINT list	PRINT list of exp	PRINT list of exp	PRINT list
PRINT #	see WRITE #	PRINT # exp, exp; list END	PRINT # N, list PRINT: N, list	PRINT # exp, list	PRINT ON num: list	see WRITE	PRINT: num; key, list
PRINT # USING	--------	--------	PRINT USING string # var, line #, list	PRINT # exp, USING; str exp, list	--------	see WRITE	--------

**Commands and elements that can be used.

7B

455

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
NOT	------	------	------	------	------	------	NOT
NOT EQUAL**	<> or ≠	<>	<>	<>	<> or ><	<> or ><	<> or #
NUMERIC VARIABLE NAME**	letter,$,@,#or letter,$,@,# digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit
ON	------	------	ON exp GOTO line#,... line#	ON exp GO TO line#,... line#	ON exp GOTO line#,... ON exp THEN line#,...	------	------
ON ERROR GOTO	------	------	------	------	------	------	------
OPEN	OPEN exp,var, INPUT OPEN exp,var, OUTPUT	------	------	------	------	------	OPEN-name,num, num OPE-name,num, num
OR	------	------	------	------	------	------	OR
PAUSE	PAUSE comment	------	------	------	------	------	------
PLOT	------	------	------	------	------	------	------
PRINT **	PRINT list	PRING list	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list
PRINT #	see WRITE #	------	see WRITE #	PRINT FILE (name) list	------	PRINT # exp list	PRINT # exp, list
PRINT # USING	------	------	------	------	------	------	------

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

7C

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC	7D
NOT	--------	--------	--------	--------	--------	--------	NOT	
NOT EQUAL **	<>	\NE or <> or >< or ≠	<>	<>	<> or ><	<>	<>	
NUMERIC VARIABLE NAME **	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	
ON	--------	ON exp GOTO line #,...	--------	--------	ON var GOTO line #,...	--------	--------	
ON ERROR GOTO	--------	--------	--------	--------	ON STATUS GO TO line#,...	--------	--------	
OPEN	--------	--------	--------	--------	OPEN (num) name	--------	--------	
OR	--------	--------	--------	--------	OR	--------	OR	
PAUSE	--------	--------	--------	--------	--------	--------	WAIT exp	
PLOT	--------	--------	--------	--------	--------	--------	--------	
PRINT **	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list	
PRINT #	--------	PRINT# exp,list PRINT FILE name, list	--------	--------	PRINT (num,IND=exp, ERR=num,END=num) list	--------	--------	
PRINT # USING	--------	--------	--------	--------	see WRITE # USING	--------	--------	

**Commands and elements that can be used.

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16
NOT	--------	--------	--------	NOT	NOT	NOT	--------
NOT EQUAL**	<> or ≠	<>	<>	<> or #	<> or >< or #	<> or >< or #	<>
NUMERIC VARIABLE NAME**	letter, @, #, $; or a letter, @, #, or $ digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit
ON	--------	ON exp GOTO line #, ...	see GOTO OF	ON exp GOTO line #, ... ON exp GOSUB line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...
ON ERROR GOTO	--------	--------	--------	ON ERROR GOTO line #	--------	--------	--------
OPEN	--------	--------	OPEN |name|, OUTPUT OPEN |name|, INPUT	OPEN |file|, options, exp	--------	--------	--------
OR	--------	--------	--------	OR	OR	OR	OR
PAUSE	PAUSE comment or PAUSE or SUSPEND	--------	--------	PAUSE	--------	--------	--------
PLOT	--------	--------	--------	--------	--------	--------	--------
PRINT**	PRINT list	PRINT list	PRINT list	PRINT list or DISPLAY or OUTPUT or TYPE	PRINT list	PRINT list	PRINT list or PRINT $ output device $ list
PRINT #	see WRITE	--------	PRINT FILE, var 1, ...	PRINT ON exp: list	--------	--------	--------
PRINT # USING	--------	--------	--------	PRINT IN FORM str exp: exp 1, ...	--------	--------	--------

7E

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
NOT	NOT(exp)	--------	--------	--------	--------	NOT	--------
NOT EQUAL**	NEQ(exp 1, exp 2) or <> or ><	<> or ><	<> or ><	<> or ><	NE or <> or ><	# or <>	<> or #
NUMERIC VARIABLE NAME**	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit
ON	ON exp GOTO line #, ... ON exp THEN line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp GOTO line #, ...	ON exp THEN line #, ... ON exp GOTO line #, ...	--------	ON exp THEN list #, ... ON exp GOTO line #, ...
ON ERROR GOTO	--------	--------	--------	--------	--------	--------	--------
OPEN	OPEN name FOR options A FILE num	see FILES	--------	see FILES	see FILES	OPEN name, num	--------
OR	IOR(exp 1, exp 2)	--------	--------	--------	--------	OR	--------
PAUSE	PAUSE or BRK	--------	--------	--------	--------	--------	--------
PLOT		--------	--------	--------	--------	--------	--------
PRINT**	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list
PRINT #	PRINT ON exp: list	see WRITE #	--------	PRINT: name: list PRINT # num, list	PRINT # num, list	PRINT # exp; list	--------
PRINT # USING	see PRINT USING	see WRITE # USING	--------	PRINT # num, USING list	PRINT # num, USING num, list	--------	--------

**Commands and elements that can be used.

7F

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200		7G
NOT	--------	--------	NOT	--------	--------	--------		
NOT EQUAL**	#	<> or ><	# or <>	<>	<>	<>		
NUMERIC VARIABLE NAME**	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit	letter or letter digit		
ON	ON exp GOTO line #, ...	--------	see GOTO OF see GOSUB OF	see GOTO OF	ON exp GOTO line #, ...	--------		
ON ERROR GOTO	--------	--------	--------	--------	--------	--------		
OPEN	--------	--------	see ASSIGN	see FILES	see FILES	see ASSIGN		
OR	--------	--------	OR	OR (exp 1, ... exp n)	--------	--------		
PAUSE	--------	--------	--------	--------	--------	--------		
PLOT	--------	--------	--------	--------	--------	--------		
PRINT**	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list	PRINT list		
PRINT #	--------	--------	PRINT # exp; list PRINT # exp, exp 1; list	--------	--------	--------		
PRINT # USING	--------	--------	--------	--------	--------	--------		

**Commands and elements that can be used.

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C	8A
PRINT USING	--------	PRINT USING ln, exp, ...	PRINT USING string var, list	PRINT USING "string", list	PRINT USING string var, list PRINT USING line #, list	--------	PRINT USING string exp; list	
RANDOMIZE	--------	--------	RANDOMIZE	RANDOM	RANDOMIZE RANDOM RAN	--------	--------	
READ **	READ varl, ...	READ varl, ...	READ varl, ...	READ varl, ...	READ varl, ...	READ varl, ...	READ varl, ...	
READ FORWARD	--------	--------	--------	--------	READ FORWARD exp, list	--------	--------	
READ #	READ FILE (name) list	--------	READ # exp: varl, ...	READ FILE [exp] varl, ... READ FILE [exp, exp] varl, ...	READ # num, list READ: num, list	READ # exp; list READ # exp, exp; list	READ # exp; list READ # exp, exp; list	
READ (,)	--------	--------	see READ #	--------	--------	see READ #	see READ #	
RELEASE	--------	--------	--------	--------	--------	--------	--------	
REM **	REM message	REM message	REM message	REM message	REM message	REM message	REM message	
RESET	--------	see RESTORE	RESET # exp: exp	--------	--------	see REWIND	see REWIND	
RESTORE **	RESTORE RESTORE FILE (name)	RESTORE	see RESET	RESTORE	RESTORE # exp RESTORE RESTORE: exp	RESTORE RESTORE line #	RESTORE RESTORE line #	
RESUME	--------	--------	--------	--------	--------	--------	--------	
RETURN **	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN	

**Commands and elements that can be used.

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
PRINT USING	PRINT USING num, list	PRINT USING num, list	PRINT USING string exp, list PRINT USING num, list	PRINT USING string exp, list	--------	PRINT USING n, list	PRINTUSING num, list
RANDOMIZE	--------	--------	RANDOM RANDOMIZE	RANDOMIZE	RANDOMIZE	RANDOMIZE	--------
READ **	READ list	READ var1, ...	READ var1, ...	READ var1, ...	READ var1, ...	READ var1, ...	READ var1, ...
READ FORWARD	--------	--------	--------	--------	--------	--------	--------
READ #	--------	READ # exp, exp; list	READ # N, list READ: N, list	see INPUT	--------	see READ ( )	GET: num; key, list
READ (,)	--------	--------	--------	see INPUT	--------	READ (60, n) list	--------
RELEASE	--------	--------	--------	see KILL	--------	RELEASE u	--------
REM **	REM message	REM message	REM message	REM message	REM message or #	REM or # message	REM message or * message
RESET	RESET name	--------	--------	see RESTORE	--------	see RESTORE, SETPTR	--------
RESTORE **	RESTORE	RESTORE RESTORE line #	RESTORE list RESTORE * RESTORE RESTORE$	RESTORE	RESTORE RESTORE * RESTORE$	RESTORE	RESTORE RESTORE line #
RESUME	--------	--------	--------	RESUME line #	--------	--------	--------
RETURN**	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN

8B

**Commands and elements that can be used.

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
PRINT USING	PRINT USING line#,expl,...	------	PRINT,line#, varl,...	------	------	------	PRINT USING string;varl,...
RANDOMIZE	------	RANDOMIZE	------	------	RANDOM RANDOMIZE	------	------
READ**	READ varl,...	READ varl,...	READ varl,...	READ varl,...	READ varl,...	READ varl,...	READ varl,...
READ FORWARD	------	------	------	------	------	------	------
READ #	GET exp: varl,...	------	READ # exp, varl,...	READ FILE (name)varl,...	------	READ # exp, varl,...	READ # exp, varl,...
READ (,)	see READ #	------	see READ #	see READ #	------	see READ #	see READ #
RELEASE	------	------	------	------	------	------	------
REM**	REM message	REM message	REM message	REM message	REM message REMARK message	REM message REMARK message	REM message
RESET	RESET expl,...	------	see REWIND	RESTORE FILE (name)	------	RESTORE # exp	READ # num, 1
RESTORE**	RESTORE Comment	RESTORE	RESTORE	RESTORE	RESTORE RESTORE * RESTORE $	RESTORE	RESTORE RESTORE num
RESUME	------	------	------	------	------	------	------
RETURN**	RETURN comment	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

	NCR CENTURY 100 BASIC I	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC	8D
PRINT USING	--------	--------	--------	--------	--------	--------	--------	
RANDOMIZE	--------	--------	--------	--------	--------	--------	--------	
READ **	READ var1,...	READ var1,...	READ var1, ...	READ var1, ...	--------	READ var1,...	READ var1....	
READ FORWARD	--------	--------	--------	--------	--------	--------	--------	
READ #	--------	see INPUT #	--------	--------	READ (num,IND=exp,ERR=num, END=num) list or EXTRACT	--------	--------	
READ (,)	--------	see INPUT #	--------	--------	see read #	--------	--------	
RELEASE	--------	--------	--------	--------	ERASE name	--------	--------	
REM **	REM message	REM message	& message REM message	& message REM message	: message REM message	REM message	REM message	
RESET	--------	RESTORE FILE name RESTORE # exp	--------	--------	--------	--------	--------	
RESTORE **	RESTORE	RESTORE	RESTORE	RESTORE	--------	--------	RESTORE	
RESUME	RESUME line #	--------	--------	--------	--------	--------	--------	
RETURN **	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN	

**Commands and elements that can be used.

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEXBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16
PRINT USING	PRINT USING line #, list	--------	SET DIGITS SET FORMAT not really formatted print	PRINT IN FORM str exp: exp 1, ...	--------	--------	--------
RANDOMIZE	--------	--------	--------	--------	--------	--------	--------
READ**	READ var 1, ...	READ var 1, ...	READ var 1, . .	READ var 1, ...	READ var 1, ...	READ var 1, ...	READ var 1, ...
READ FORWARD	--------	--------	--------	--------	--------	--------	--------
READ #	GET name, var, var 1, ...	READ # exp, var 1, var 2, ...	READ FILE var 1, ...	see INPUT FROM	--------	--------	--------
READ (,)	see READ #	see READ #	see READ #	see INPUT FROM	--------	--------	--------
RELEASE	--------	--------	--------	see INPUT FROM	--------	--------	--------
REM**	REM comment	REM comment	REM comment or ! comment	REM comment or ! comment	REM comment	REM comment	REM comment
RESET	RESET name, ...	see REWIND	--------	--------	--------	--------	--------
RESTORE**	RESTORE comment or RESTORE	RESTORE	RESTORE RESTORE * RESTORE $	RESTORE	RESTORE	RESTORE	RE TORE
RESUME	RESUME or GO	--------	PROCEED AGAIN	--------	--------	--------	--------
RETURN**	RETURN comment or RETURN	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN

8E

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
PRINT USING	PRINT ON exp IN FORM IMAGE str: var 1, ...	PRINT, num, list	--------	PRINT USING num, list	PRINT USING num, list	--------	--------
RANDOMIZE	RANDOMIZE	--------	--------	--------	--------	--------	--------
READ**	READ var 1, ..., var n	READ var 1, ..., var n	READ var 1, ..., var n	READ var 1, ... var n	READ var 1, ... var n	READ var 1, ... var n	READ var 1, .. var n
READ FORWARD	--------	--------	--------	--------	--------	READ # exp, exp	--------
READ #	see INPUT FROM	READ # exp, var 1, var 2, ...	--------	READ # num, var 1, ...	READ # num, list READ: num, list	READ # exp; var 1, ... var n	--------
READ (,)	see INPUT FROM	see READ #	--------	see READ #	see READ #	see READ #	--------
RELEASE	--------	--------	--------	--------	--------	--------	--------
REM**	REM comment may follow statements after special character	REM comment or * comment	REM comment	REM comment	REM comment	REM comment	REM comment may follow statements after special character
RESET	--------	--------	--------	see REWIND	see REWIND	READ # exp, 1	--------
RESTORE**	RESTORE or RESTORE * or RESTORE $	RESTORE	RESTORE	RESTORE RESTORE * RESTORE $	RESTORE RESTORE * RESTORE $	RESTORE RESTORE line #	RESTORE RESTORE * RESTORE $
RESUME	--------	--------	--------	--------	--------	--------	--------
RETURN**	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN	RETURN

**Commands and elements that can be used.

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200	
PRINT USING	--------	--------	PRINT USING num; list PRINT USING str var; list PRINT USING str; list	PRINT USING num, list	PRINT USING num, list	PRINT USING num, list	
RANDOMIZE	RANDOMIZE	--------	--------	RANDOM	--------	--------	
READ**	READ var 1, ... var n	READ var 1, ... var n	READ var 1, ... var n	READ var 1, ... var n	READ var 1, ... var n	READ var 1, var n	
READ FORWARD	--------	--------	ADVANCE # exp; exp, var	--------	--------	--------	
READ #	--------	--------	READ # exp; var 1, ... var n READ # exp, exp; var 1, ..., var n	FILE READ # num, var 1, ..., var n	READ # exp, var 2, ... var n READ: exp, var 1, ..., var n	DATALOAD	
READ (,)	--------	--------	see READ #	see READ #	see READ #	--------	
RELEASE	--------	--------	--------	--------	--------	--------	
REM**	REM comment	REM comment	REM comment	REM comment	REM comment	REM comment	
RESET	--------	--------	--------	--------	see REWIND	--------	
RESTORE**	RESTORE	RESTORE	RESTORE RESTORE num	RESTORE RESTORE num	RESTORE	RESTORE RESTORE num	
RESUME	--------	RESUME	--------	--------	--------	--------	
RETURN**	RETURN	RETURN	RETURN RETURN exp	RETURN	RETURN	RETURN	

**Commands and elements that can be used.

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C
REWIND	--------	--------	--------	--------	--------	READ # exp, I	READ # exp, I
SCRATCH	--------	--------	SCRATCH # exp	--------	SCRATCH # exp SCRATCH: exp	--------	--------
SETPTR	--------	--------	--------	--------	SETW exp TO exp	see READ #	see READ #
SMALLEST #**	E-368	5.4 E-79	1.46937 E-39	5.4E-79	1.46937 E-39	E-38	E-38
STOP**	STOP	STOP	STOP	STOP	STOP	STOP	STOP
STRING QUOTES **	"	'	"	"	"	"	"
STRINGS	--------	--------	--------	--------	--------	--------	--------
STRING VARIABLE NAME **	letter $	letter $	Num name followed by $	letter $	letter $	letter $	letter $
SUB	--------	--------	SUB name: arg list	--------	--------	--------	--------
SUBEND	--------	--------	SUBEND	--------	--------	--------	--------
TIME	--------	--------	TIME n	--------	--------	--------	--------
UNLESS	--------	--------	--------	--------	--------	--------	--------
UNTIL	--------	--------	--------	--------	--------	--------	--------
USER DEFINED FUNCTION NAMES**	FN letter	FN letter	FN letter	FN letter	FN letter	FN letter	FN letter

**Commands and elements that can be used.

9A

463

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2,0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
REWIND	--------	--------	--------	--------	--------	REWIND num	--------
SCRATCH	--------	--------	SCRATCH list	see KILL	--------	SCRATCH num	--------
SETPTR	--------	--------	SET N, exp ...	--------	--------	SETPTR num1, var	--------
SMALLEST #**	5.4 E-79	E-38	2.4 E-39	.14 E-38	E-39	E-99	5.398 E-79
STOP**	STOP or STOP message	STOP	STOP	STOP	STOP	STOP	STOP
STRING QUOTES**	' or "	"	"	"	'	"	'
STRINGS	--------	--------	--------	--------	STRINGS num default 60	--------	--------
STRING VARIABLE NAME**	alphabetic $	letter $	Numeric name $	Numeric name $	Numeric name $	Numeric name $	letter $
SUB	--------	--------	--------	--------	--------	--------	--------
SUBEND	--------	--------	--------	--------	--------	--------	--------
TIME	--------	--------	--------	--------	--------	--------	--------
UNLESS	--------	--------	--------	UNLESS condition	--------	--------	--------
UNTIL	--------	--------	--------	UNTIL condition	--------	--------	--------
USER DEFINED FUNCTION NAMES **	FN letter	FN letter	FN letter	FN followed by variable	FN letter	FN letter	FN letter

**Commands and elements that can be used.

9B

403

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
REWIND	see RESET	------	RESTORE # exp	see RESET	------	see RESET	see RESET
SCRATCH	------	------	SCRATCH # exp	------	------	------	------
SETPTR	------	------	------	------	------	------	------
SMALLEST #**	1E-78	1E-615	1E-616	1E-368	5.4E-79	1E-47	1E-38
STOP**	STOP comment	STOP	STOP	STOP	STOP	STOP	STOP
STRING QUOTES**	" or '	------	"	"	"	"	"
STRINGS	------	------	------	------	------	------	------
STRING VARIABLE NAME**	letter $	------	Numeric name $	letter $	Numeric name $	Numeric name $	letter $
SUB	------	------	------	------	------	------	------
SUBEND	------	------	------	------	------	------	------
TIME	------	------	------	------	------	------	------
UNLESS	------	------	------	------	------	------	------
UNTIL	------	------	------	------	------	------	------
USER DEFINED FUNCTION NAMES **	FN letter	FN letter	FN letter	FN letter	------	FN letter	FN letter

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700
**Commands and elements that can be used.

9C

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUGHS B250C BASIC	BURROUGHS B3500 BASIC	BASIC FOUR. BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC
REWIND	--------	--------	--------	--------	--------	--------	--------
SCRATCH	--------	--------	--------	--------	ERASE name	--------	--------
SETPTR	--------	--------	--------	--------	REKEY	--------	--------
SMALLEST # **	1E-99	8.758E-47	1E-99	1E-99	1E-99	1.67E-57	1E-99
STOP **	STOP	STOP	STOP	STOP	STOP	--------	STOP
STRING QUOTES **	"	"	"	"	"	"	"
STRINGS	--------	--------	--------	--------	--------	--------	--------
STRING VARIABLE NAME **	--------	letter $	--------	--------	num name $	--------	--------
SUB	--------	--------	--------	--------	--------	--------	--------
SUBEND	--------	--------	--------	--------	--------	--------	--------
TIME	--------	--------	--------	--------	--------	--------	--------
UNLESS	--------	--------	--------	--------	--------	--------	--------
UNTIL	--------	--------	--------	--------	--------	--------	--------
USER DEFINED FUNCTION NAMES **	FN letter	FN letter	FN letter	FN letter	FN letter	--------	FN letter

**Commands and elements that can be used.

9D

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16	9E
REWIND	see RESET	RESTORE # exp	--------	--------	--------	--------	--------	
SCRATCH	--------	SCRATCH # exp	SCRATCH	--------	--------	--------	--------	
SETPTR	--------	--------	--------	--------	--------	--------	--------	
SMALLEST #**	1E-99	5.78960E-76	5E-76	5E-76	2.71E-20	2.71E-20	2.71E-20	
STOP**	STOP comment or STOP	STOP	STOP	STOP	STOP	STOP	STOP	
STRING QUOTES**	'	"	"	" or '	"	"	"	
STRINGS	--------	--------	--------	--------	--------	--------	--------	
STRING VARIABLE NAME**	letter, $, @, or # followed by $	letter $	num name $	num name $	--------	letter $	letter $	
SUB	--------	--------	--------	--------	--------	--------	--------	
SUBEND	--------	--------	--------	--------	--------	--------	--------	
TIME	--------	--------	--------	TIME	--------	--------	--------	
UNLESS	--------	--------	--------	--------	--------	--------	--------	
UNTIL	--------	--------	--------	UNTIL exp	--------	--------	--------	
USER DEFINED FUNCTION NAMES**	FN letter, $, @, or #	FN letter	FN letter	FN letter	FN letter	FN letter	FN letter	

**Commands and elements that can be used.

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
REWIND	--------	RESTORE # exp	--------	RESTORE # num	RESTORE # num RESTORE: num	--------	--------
SCRATCH	--------	--------	--------	SCRATCH :name: SCRATCH # num	SCRATCH # num SCRATCH: num	--------	--------
SETPTR	--------	--------	--------	--------	SET: num TO exp	--------	--------
SMALLEST #**	1E-39	1E-38	1E-38	5.7896E-76	1E-38	1E-38	1E-39
STOP**	STOP	STOP	STOP	STOP	STOP	STOP	STOP
STRING QUOTES**	"	"	"	"	"	"	"
STRINGS	STRING num	--------	--------	--------	--------	--------	--------
STRING VARIABLE NAME**	num name $	letter $	--------	letter $	num name $	letter $	letter $
SUB	--------	--------	--------	--------	--------	--------	--------
SUBEND	--------	--------	--------	--------	--------	--------	--------
TIME	--------	--------	--------	--------	--------	--------	--------
UNLESS	--------	--------	--------	--------	--------	--------	--------
UNTIL	UNTIL condition	--------	--------	--------	--------	--------	--------
USER DEFINED FUNCTION NAMES**	FN letter	FN letter	FN letter	FN letter	FN letter	FN letter	FN letter

**Commands and elements that can be used.

9F

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200	
REWIND	--------	--------	--------	--------	RESTORE # exp RESTORE: exp	--------	
SCRATCH	--------	--------	PURGE name	FILEMOD # exp, option	SCRATCH # exp SCRATCH: exp	--------	
SETPTR	--------	--------	--------	--------	SET: exp, var	--------	
SMALLEST #**	1E-37	1E-99	1E-77	1E-65	4.31809E-78	1E-100	
STOP**	STOP	STOP	STOP	STOP	STOP	STOP STOP "comment" STOP digit	
STRING QUOTES**	"	"	"	"	"	"	
STRINGS	--------	--------	--------	--------	--------	--------	
STRING VARIABLE NAME**	--------	--------	letter $ or letter digit $	letter $	letter $	num name $	
SUB	--------	--------	--------	--------	--------	DEFFN' num (var 1, ... var n) DEFFN' (string)	
SUBEND	--------	--------	--------	--------	--------	--------	
TIME	--------	--------	--------	--------	--------	--------	
UNLESS	--------	--------	--------	--------	--------	--------	
UNTIL	--------	--------	--------	--------	--------	--------	
USER DEFINED FUNCTION NAMES**	FN letter	FN letter	FN letter	FN letter	FN letter	FN letter	

**Commands and elements that can be used.

9G

474

10A

	BASIC 2.0 CDC 6600 SCOPE	IBM CPS UNIV OF IOWA	DARTMOUTH	DATA GENERAL	GE MARK II GE MARK III	HP2000B	HP2000C
WHILE	--------	--------	--------	--------	--------	--------	--------
WRITE #	WRITE FILE (name) list	--------	WRITE # exp: list	WRITE FILE [exp] list or WRITE FILE [exp, exp] list	WRITE # exp, list WRITE: exp, list	see PRINT #	see PRINT #
WRITE # USING	--------	--------	--------	--------	WRITE # exp USING str, list	--------	--------
WRITE (,)	--------	--------	see WRITE #	--------	--------	see PRINT #	see PRINT #

	IBM ITF	LEASCO	PDP 10	PDP 11	UNIVAC 1100 UBASIC VERSION 2.0 MANKATO STATE CLG	MULTICOMP OR UNIV MASS BASICX	XEROX
WHILE	--------	--------	--------	WHILE condition	--------	--------	--------
WRITE #	PUT 'name', list	--------	WRITE # N, list WRITE: N, list	see PRINT	WRITE ON num: list	see WRITE (,)	PUT: num, key, list
WRITE # USING	--------	--------	--------	--------	--------	--------	--------
WRITE (,)	--------	--------	--------	see PRINT	--------	WRITE (μ, n) list	--------

10B

10C

	IBM CALL/360-OS	PDP 8/E	HONEYWELL 200	CDC 6000 KRONOS BASIC 2.0	NCR CENTURY 200	UCSD* BASIC B6700	HP2000F
WHILE	------	------	------	------	------	------	------
WRITE #	PUT exp: varI,...	------	WRITE # exp, varI,...	WRITE FILE (name)expI,...	------	see PRINT #	see PRINT #
WRITE # USING	------	------	------	------	------	------	------
WRITE (,)	see WRITE #	------	see WRITE #	see WRITE #	------	see PRINT #	see PRINT #

*UNIVERSITY OF CALIFORNIA, SAN DIEGO, BURROUGHS B6700

	NCR CENTURY 100 BASIC 1	BURROUGHS B5500 BASIC	BURROUGHS B2500 BASIC	BURROUGHS B3500 BASIC	BASIC FOUR BUSINESS BASIC	UNICOMP COMP 16 or COMP 18 BASIC	VARIAN 620 or V73 BASIC
WHILE	--------	--------	--------	--------	--------	--------	--------
WRITE #	--------	see PRINT #	--------	--------	WRITE (num, IND=exp,ERR= num,END=num) list	--------	--------
WRITE # USING	--------	--------	--------	--------	WRITE USING num (num,IND=exp,ERR= num,END=num) list	--------	--------
WRITE (,)	--------	see PRINT #	--------	--------	see WRITE #	--------	--------

10D

	IBM S3 MOD 6 BASIC	GE 255 TIME SHARING BASIC	COM-SHARE BASIC	COM-SHARE NEWBASIC	WESTINGHOUSE BASIC II	WESTINGHOUSE BASIC III	GENERAL AUTOMATION ADVANCED BASIC-16	10E
WHILE	--------	--------	--------	WHILE exp	--------	--------	--------	
WRITE #	PUT name, var 1, var 2, ...	WRITE # exp, var 1, var 2, ...	see PRINT #	see PRINT FROM	--------	--------	--------	
WRITE # USING	--------	--------	--------	--------	--------	--------	--------	
WRITE (,)	see WRITE #	see WRITE #	see PRINT #	see PRINT FROM	--------	--------	--------	

	UNIVAC 1100 UBASIC	HONEYWELL 1640 XBASIC	HONEYWELL 316, 516, and 716 BASIC	HONEYWELL 400 XBASIC	HONEYWELL 600 BASIC	HP2000E	UNIVAC 1100 UNIV OF MARYLAND RELEASE V 1.3
WHILE	WHILE condition	--------	--------	--------	--------	--------	--------
WRITE #	WRITE ON exp: list	WRITE # exp, list	--------	WRITE # num, list	WRITE # num, list WRITE: num, list	--------	--------
WRITE # USING	see WRITE #	WRITE # exp, num, list	--------	--------	--------	--------	--------
WRITE (,)	see WRITE #	see WRITE #	--------	--------	see WRITE #	--------	--------

10F

480

	MICRODATA BASIC	Q-DATA BASIC-1	HP3000	WANG 3300	GENERAL ELECTRIC MARK I	WANG 2200	
WHILE	--------	--------	--------	--------	--------	--------	
WRITE #	--------	--------	see PRINT #	FILEWRITE # num, list	WRITE # exp, var 1, ... var n WRITE : exp, var 1, ... var n	DATASAVE	
WRITE # USING	--------	--------	--------	--------	--------	--------	
WRITE (,)	--------	--------	see PRINT #	see WRITE #	see WRITE #	see WRITE #	

The following 4 tables are a list of the BASIC built-in functions where:

ABS: = Absolute value
ACS: = Ascii
ASN: = Arcsin
ATN: = Arctangent
BOOL: = Returns true value of relation
CLK: = Time of day
COL: = Next print position
COS: = Cosine
COT: = Cotangent
CSC: = Cosecant
CSF: = Returned statuscode
DAT: = Date
DEG: = Degrees from radians
DET: = Determinent
DIG: = Digital part from scientific notation
DIV: = Integer division
EOF: = End of file
EPT: = Exponent part
EXP: = Exponentiation
FIX: = Truncation
FLD: = Selects bits
FRP: = Fractional part
GET: = Field data equivalent
HCS: = Hyperbolic cosine
HTN: = Hyperbolic tangent
INP: = Integer part
INS: = Converts to binary integer
INT: = Largest integer
KEY: = Next available position of file

LIN: = Skips lines
LOC: = Location of file pointer
LOF: = Length of file
LOG: = Natural logarithms
LOG10: = Common logarithms
LTW: = Logarithm base 2
MAR: = Margin for file
MOD: = X - Y*INT(X/Y)
MXL: = Maximum length of string
NUM: = Number of data input
PI: = 3.1415927
PIX: = π times argument
POS: = Location of string
RAD: = Radians
REKEY: = Change position number of record in file
RND: = Random number
RUN: = Elapsed time
SEC: = Secant
SGN: = Algebraic sign
SIN: = Sine
SPA: = Skips spaces
SPC: = Outputs a number of spaces
SQR: = Square root
TAB: = Tabulation
TAN: = Tangent
TIM: = Elapsed time
TIS: = Time of day in milliseconds
TYP: = Type of file
XPT: = Exponent part

	ABS	ACS	ASN	ATN	BOOL	CLK	COL	COS	COT	CSC	CSF	DAT	DEG	DET	DIG	DIV	EOF	EPT	EXP	FIX	FLD	FRP	GET	HCS	HSN	HTN	INP	INS	INT	KEY
BASIC 2.0 CDC 6600 SCOPE	X			X		X		X											X										X	
IBM CPS UNIV OF IOWA	X			X				X											X										X	
DARTMOUTH	X			X				X	X					X					X										X	
DATA GENERAL	X			X				X						X					X										X	
GE MARK II & III	X			X				X	X					X					X										X	
HP2000B	X			X				X											X										X	
HP2000C	X			X				X											X										X	
IBM ITF	X	X	X	X				X	X	X			X	X					X					X	X	X			X	
LEASCO	X			X				X											X										X	
PDP 10	X			X				X	X					X					X										X	
PDP 11	X			X				X											X	X									X	
UNIVAC 1100 VER 2.0 MANKATO STATE CLG	X			X			X	X	X					X					X			X	X				X		X	
MULTICOMP OR UNIV MASS BASICX	X			X				X						X			X		X										X	
XEROX	X	X	X	X				X	X	X			X						X					X	X	X			X	
IBM CALL/360-OS	X	X	X	X				X	X	X			X						X					X	X	X			X	
PDP 8/E	X			X				X											X										X	
HONEYWELL 200	X			X		X		X											X										X	
CDC 6000 KRONOS BASIC 2.0	X			X		X		X											X										X	
NCR CENTURY 200	X			X				X	X										X										X	
UCSD BASIC B6700	X			X				X											X										X	
HP2000F	X			X				X											X										X	
NCR CENTURY 100	X			X				X											X										X	
BURROUGHS-B5500	X			X				X								X			X										X	
BURROUGHS-B2500	X			ATAN				X											X										X	

	ABS	ACS	ASN	ATN	BOOL	CLK	COL	COS	COT	CSC	CSF	DAT	DEG	DET	DIG	DIV	EOF	EPT	EXP	FIX	FLD	FRP	GET	HCS	HSN	HTN	INP	INS	INT	KEY
BURROUGHS-B3500	X			ATAN				X											X										X	
BASIC-4 BUSINESS BASIC	X																	X				FPT							X	X
UNICOMP-BASIC				X				X											X											
VARIAN-620, V73	X			X				X											X										X	
IBM S3 MOD 6	X	X	X	X				X	X	X			X						X					X	X	X			X	
GE 235 TIME SHARING	X			X		X		X											X										X	
COM-SHARE BASIC	X			X				X											X										X	
COM-SHARE NEWBASIC	X	ARCCOS	ARCSIN	ATAN				X											X	X				COSH	SINH	TANH			X	
WESTINGHOUSE BASIC II	X																		X										X	
WESTINGHOUSE BASIC III	X			X				X											X										X	
GENERAL AUTOMATION BASIC-16 ADVANCED	X			X				X											X										X	
UNIVAC 1100 UBASIC VERSION 3.2	X			X		X	X	X	X		X	X		X	X				X		X	X					X	X	X	
HONEYWELL 1640 XBASIC	X			X				X											X										X	
HONEYWELL 316, 516, 716	X			X				X											X										X	
HONEYWELL 400 XBASIC	X			X				X											X										X	
HONEYWELL 600 BASIC	X			X				X	X					X					X										X	
HP2000E	X			X				X											X										X	
UNIVAC 1100 UNIV OF MARYLAND	X			X				X	X					X					X								X		X	
MICRODATA	X			X				X											X										X	
Q-DATA BASIC-1	X			X				X											X										X	
HP3000	X			X				X											X					CSH	SNH	TNH			X	
WANG 3300	X			X	X			X											X										X	
GE MARK I	X			X		X		X											X										X	
WANG 2200	X	ARCCOS	ARCSIN	X				X											X										X	

434

	LIN	LOC	LOF	LOG	LOG10	LTW	MAR	MOD	MXL	NUM	PI	PIX	POS	RAD	REKEY	RND	RUN	SEC	SGN	SIN	SPA	SPC	SQR	TAB	TAN	TIM	TIS	TYP	XPT
BASIC 2.0 CDC 6600 SCOPE				X												X			X	X			X		X	X			
IBM CPS UNIV OF IOWA				X	ALGT											X			X	X			X						
DARTMOUTH		X	X	X			X	X		X						X			X	X			X		X	X			
DATA GENERAL				X												X			X	X			X		X				
GE MARK II & III				X												X			X	X			X	X	X	X			
HP2000B				X												X			X	X			X	X	X	X		X	
HP2000C				X												X			X	X			X	X	X	X			
IBM ITF				X	LGT	X										X		X	X	X			X		X				
LEASCO				X												X			X	X			X	X	X	X		X	
PDP 10		X	X	X						X						X			X	X			X	X	X				
PDP 11				X	X						X		X			X			X	X			X	X	X				
UNIVAC 1100 VER 2.0 MANKATO STATE CLG				X	LGT			X		X						X			X	X			X	X	X	TIS		X	
MULTICOMP OR UNIV MASS BASICX				X						X						X	X		X	X			X	X	X				
XEROX				X	LGT	X										X		X	X	X			X	X	X	X			
IBM CALL/360-OS				X	LGT		X				X			X		X		X	X	X			X		X				
PDP 8/E				X												X			X	X			X	X	X				
HONEYWELL 200				X												X			X	X			X		X	X			
CDC 6000 KRONOS BASIC 2.0				X												X			X	X			X		X	X			
NCR CENTURY 200				X	COM											X			X	X			X	X	X				
UCSD BASIC B6700				X												X			X	X			X	X	X				
HP2000F				X												X			X	X			X	X	X				
NCR CENTURY 100				X												X			X	X			X		X				
BURROUGHS-B5500				X				X								X			X	X			X	X	X	X		X	
BURROUGHS-B2500				X												X			X	X			X	X	X				

	LIN	LOC	LOF	LOG	LOG10	LTW	MAR	MOD	MXL	NUM	PI	PTX	POS	RAD	REKEY	RND	RUN	SEC	SGN	SIN	SPA	SPC	SQR	TAB	TAN	TIM	TIS	TYP	XPT
BURROUGHS-B3500				X												X			X	X			X	X	X				
BASIC-4 BUSINESS BASIC															X				X										
UNICOMP-BASIC				X																X			X						
VARIAN-620, V73				X												X			X	X			X	X	X				
IBM S3 MOD 6				X	LGT	X					SPI			X		X		X	X	X			X		X				
GE 255 TIME SHARING				X												X			X	X			X	X	X	X			
COM-SHARE BASIC				X	LGT						X					X			X	X			X	X	X				
COM-SHARE NEWBASIC				X	X			X			X		X			NUM			X	X			X		X	TIME			
WESTINGHOUSE BASIC II				X				X											X					X					
WESTINGHOUSE BASIC III				X				X								X			X	X			X	X	X				
GENERAL AUTOMATION BASIC-16 ADVANCED				X				X								X			X	X			X	X	X				
UNIVAC 1100 UBASIC VERSION 3.2				X	LGT				X	X						X			X	X			X	X	X	X	X		X
HONEYWELL 1640 XBASIC				X												X			X	X			X	X	X	X			
HONEYWELL 316, 516, 716				X												X			X	X			X	X	X				
HONEYWELL 400 XBASIC				X	CLG											X			X	X			X	X	X	X			
HONEYWELL 600 BASIC				X	CLG											X			X	X		X	X*	X	X	X			
HP2000E				X												X			X	X			X	X	X	X		X	
UNIVAC 1100 UNIV OF MARYLAND				X	LGT			X		X						X			X	X			X	X	X				
MICRODATA				X												X			X	X			X	X	X				
Q-DATA BASIC-I				X												X			X	X			X		X				
HP3000	X			X				X				X				X			X	X	X		X	X	X	X			
WANG 3300				X												X			X	X			X	X	X				
GE MARK I				X												X			X	X			X	X	X	X			
WANG 2200				X							#PI					X			X	X			X	X	X				

The following table refers to matrix operations and built-in functions where:

CON: = Matrix of all ones

IDN: = Identity matrix

INV: = Inverse

NUL$: = Matrix of null strings

TRN: = Transpose

ZER: = Zero matrix

	A+B	A-B	A*B	K*A	CON	IDN	INV	NUL$	TRN	ZER	A*B	INPUT	PRINT	PRINT USING	READ	WRITE	GET	PUT	DIM
BASIC 2.0 CDC 6600 SCOPE	X	X	X	X	X	X	X		X	X	X	X	X		X	X			
IBM CPS UNIV OF IOWA	X	X	X	X	X	X			X	X	X				X	X			
DARTMOUTH	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			
DATA GENERAL	X	X	X	X	X	X	X		X	X	X	X	X		X				
GE MARK II & III	X	X	X	X	X	X	X		X	X	X	X	X		X				
HP2000B	X	X	X	X	X	X	X		X	X	X	X	X		X				
HP2000C	X	X	X	X	X	X	X		X	X	X	X	X		X				
IBM ITF	X	X	X	X	X	X	X		X	X	X	X	X	X	X		X	X	
LEASCO	X	X	X	X	X	X	X		X	X	X	X	X		X				
PDP 10	X	X	X	X	X	X	X		X	X	X	X	X		X				
PDP 11	X	X	X	X	X	X	X		X	X		X	X		X	X			
UNIVAC 1100 VER 2.0 MANKATO STATE CLG	X	X	X	X	X	X	X		X	X	X	X	X		X				
MULTICOMP OR UNIV MASS BASICX	X	X	X	X	X	X	X		X	X		X	X		X	X			
XEROX	X	X	X	X	X	X	X		X	X	X	X	X		X		X	X	
IBM CALL/360-OS	X	X	X	X	X	X	X		X	X	X		X		X		X	X	
PDP 8/E																			
HONEYWELL 200	X	X	X	X	X	X	X		X	X	X		X		X				
CDC 6000 KRONOS BASIC 2.0	X	X	X	X	X	X	X		X	X	X	X	X		X	X	INPUT FILE	PRINT FILE	
NCR CENTURY 200																			
UCSD BASIC B6700																			
HP2000F	X	X	X	X	X	X	X		X	X	X	X	X	X	X				
NCR CENTURY 100																			
BURROUGHS-B5500	X	X	X	X	X	X	X		X	X	X	X	X		X				
BURROUGHS-B2500	X	X	X	X	X	X	X		X	X	X	X	X		X				

	A+B	A-B	A*B	X*A	CON	IDN	INV	NUL$	TRN	ZER	A=B	INPUT	PRINT	PRINT USING	READ	WRITE	GET	PUT	DIM
BURROUGHS-B3500	X	X	X	X	X	X	X		X	X	X	X	X		X				
BASIC-4 BUSINESS BASIC																			
UNICOMP-BASIC																			
VARIAN-620, V73	X	X	X	X	X	X	X		X	X	X		X						
IBM S3 MOD 6	X	X	X	X	X	X	X		X	X	X	X	X	X	X		X	X	
GE 255 TIME SHARING	X	X	X	X	X	X	X		X	X	X		X		X				
COM-SHARE BASIC	X	X	X	X	X	X	X		X	X	X		X		X	X			
COM-SHARE NEWBASIC	X	X	X	X	X	X				X	X	X	X		X				
WESTINGHOUSE BASIC II																			
WESTINGHOUSE BASIC III	X	X	X	X	X	X	X		X	X	X		X		X				
GENERAL AUTOMATION BASIC-16 ADVANCED	X	X	X	X	X	X	X		X	X	X	X	X		X				
UNIVAC 1100 UBASIC VERSION 3.2	X	X	X	X	X	X	X		X	X	X	X	X		X				X
HONEYWELL 1640 XBASIC	X	X	X	X	X	X	X		X	X	X	X	X		X	X			
HONEYWELL 316, 516, 716																			
HONEYWELL 400 XBASIC	X	X	X	X	X	X	X		X	X	X	X	X		X				
HONEYWELL 600 BASIC	X	X	X	X	X	X	X		X	X	X	X	X		X				
HP2000E	X	X	X	X	X	X	X		X	X	X	X	X		X				
UNIVAC 1100 UNIV OF MARYLAND	X	X	X	X			X			X		X	X		X				
MICRODATA																			
Q-DATA BASIC-1																			
HP3000	X	X	X	X	X	X	X	NUL	X	X	X	X	X	X	X				
WANG 3300	X	X	X	X	X	X	X		X	X	X	X	X	X	X		FILE READ	FILE WRITE	
GE MARK I	X	X	X	X	X	X	X		X	X	X		X		X				
WANG 2200																			

The next table shows various string built-in functions. These functions are sometimes quite complicated and their descriptions should be referenced in the appropriate manual.

	CHR$(X)	ASCII(A$)	LEFT(A$,N)	RIGHT(A$,N)	MID(A$,N,M)	LEN(A$)	CLK$	DAT$	PER	POS	STR	SEG	TYP	VAL	CNT	DTS$	PAD$	TRM$	EXT$	CPY$	ADD$	SPACE$	INSTR$
BASIC 2.0 CDC 6600 SCOPE																							
IBM CPS UNIV OF IOWA																							
DARTMOUTH	X					X	X	X	X	X	X	X	X	X									
DATA GENERAL						X																	
GE MARK II & III		ASC				X	X	X			STR$			X									
HP2000B						X																	
HP2000C						X																	
IBM ITF																							
LEASCO						X																	
PDP 10	X	ASC	LEFT$	RIGHT$	MID$	X																X	X
PDP 11	X	X	X	X	X	X					STR$			X									
UNIVAC 1100 VER 2.0 MANKATO STATE CLG											STR$				X	X	X	X	X	X	X		
MULTICOMP OR UNIV MASS BASICX																							
XEROX						X					X			X									
IBM CALL/360-OS																							
PDP 8/E																							
HONEYWELL 200																							
CDC 6000 KRONOS BASIC 2.0																							
NCR CENTURY 200																							
UCSD BASIC B6700	CHR					LNG					X	X		NUM									
HP2000F						X																	
NCR CENTURY 100																							
BURROUGHS-B5500																							
BURROUGHS-B2500																							

	CHR$(X)	ASCII(A$)	LEFT(A$,N)	RIGHT(A$,N)	MID(A$,N,M)	LEN(A$)	CLK$	DAT$	PER	POS	STP	SEG	TYP	VAL	CNT	DTS$	PAD$	TRM$	EXT$	CPY$	ADD$	SPACE$	INSTR$
BURROUGHS-B3500																							
BASIC-4 BUSINESS BASIC	CHR	ASC				X								NUM									
UNICOMP-BASIC																							
VARIAN-620, V73																							
IBM S3 MOD 6											X												
GE 255 TIME SHARING																							
COM-SHARE BASIC																							
COM-SHARE NEWBASIC	CHAR	X	X	X	SUBSTR	X					X			X									
WESTINGHOUSE BASIC II																							
WESTINGHOUSE BASIC III						X																	
GENERAL AUTOMATION BASIC-16 ADVANCED						X																	
UNIVAC 1100 UBASIC VERSION 3.2						X				SEP SER				X		X	X	X	X	X	X		
HONEYWELL 1640 XBASIC																							
HONEYWELL 316,516,716																							
HONEYWELL 400 XBASIC		ASC				X	X	X		HPS													
HONEYWELL 600 BASIC		ASC			SST	X	X	X		SPC HPS	X			X									
HP2000E						X																	
UNIVAC 1100 UNIV OF MARYLAND																							
MICRODATA																							
Q-DATA BASIC-1																							
HP3000	X					X		X		X			X	NUM									
WANG 3300											X												
GE MARK I																							
WANG 2200						X					X												

COM-SHARE NEWBASIC extensions

1. LET VAR = ZERO--zeros all variables.
2. Statements may contain up to 256 characters.
3. Complex variables
4. Data type declares--INTEGER, DOUBLE INTEGER, COMPLEX, REAL, DOUBLE REAL, STRING, TEXT
5. Very much less than <<
6. Very much greater than >>
7. Binary operators--BAN conjunction, BOR disjunction, BEX exclusive or
8. Logial operation--BUT
9. Allows mixed data types and converts.
10. Comments may be added after any statement.
11. Suffix modifiers may be added after any non declarative statement.
12. Keywords may be abbreviated.
13. LET var = exp 1 = exp 2
14. NORMAL MODE IS
15. DIM var (exp: exp) as in ALGOL
16. LINK saves variables, LOAD does not.
17. May LINK or LOAD BINARY
18. APPEND in execute mode
19. FOR var = exp 1, exp 2, ...
    FOR ... UNTIL or WHILE
20. May use brackets [ ]
21. ON ESCAPE GOTO line #
22. ERASE exp FROM exp TO exp deletes material on random file
23. Formatted input--INPUT IN FORM, INPUT IN FORM FROM
24. Setting BASE
25. Suffix modifiers FOR, IF, UNLESS, UNTIL, WHILE

26. String functions: IEQIV--searches for substrings
    CTI--character to integer
    ITC--integer to character
    SPACE--returns spaces
    LITRIM--removes leading blanks
    TRIM--removes trailing blanks
    INDEX--returns position of substring
27. Functions: DIF--positive difference
    FLOAT--floating point of integer
    SNGL--single precision from double
    LSH--left shift
    RSH--right shift
    IMAG--imaginary part of complex number
    REAL--real part of complex number
    COMPLX--complex number
    CONJG--conjugate
    WAIT--halts for time
    PASS--number of times statement is executed
    REPASS--resets PASS
    DATE--12 character date
    TEL--tells if terminal buffer empty
    SIZE--length of file in words
28. Catalyst functions

Extensions of HP3000 BASIC

1. Continuation of statements by placing & as last character
2. Double precision variables
3. Complex variables
4. Integer variables
5. TYPE statements INTEGER COMPLEX LONG REAL
6. Redimensioning by REDIM
7. IF - DO and DOEND pairs
8. For loops in READ statements
9. May save extra INPUT's in a buffer and BUF function
10. Complex functions CEI, CPX, REA, IMG, CNJ
11. String functions WRD, UPS, DEB
12. Matrix functions ROW, COL
13. Functions UND, CPU, REC
14. May define type functions
15. Call external procedure in other libraries by EXT
16. UPDATE allows file to be modified.

ACT TECHNICAL BULLETIN NO. 12

THE CADA MONITOR

by

David E. Christ

The University of Iowa

The Research and Development Division

The American College Testing Program

P. O. Box 168, Iowa City, Iowa 52240

January, 1973

THE CADA MONITOR

David E. Christ

The University of Iowa

Several elements go into a Bayesian statistical analysis. Some are skilled tasks requiring the expertise of a professional and others are purely mechanical. The former include such tasks as choice of model, specification of the prior, and interpretation of the posterior distribution; whereas the latter include such things as the arithmetic necessary to take statements about the prior and combine them with the data to produce the posterior distribution and to produce probability statements about parameters using the posterior distribution. Unfortunately, it is all too often the case that the arithmetic gets in the way of the professional's decision-making task by breaking concentration and line of thought; and at times the sheer bulk of computation precludes the use of advanced techniques by the unaided researcher. For these and other reasons, a system of Computer-Assisted Data Analysis (Novick, 1971) was developed at The University of Iowa. Further investigation into available computer technology coupled with expansion of the theoretical base on which the original system rested has resulted in the refinement and expansion of the available programs and the construction of a monitor to facilitate their use.

Since CADA (Computer-Assisted Data Analysis) was meant as a research tool for general application, a search was made to find the most effective means of facilitating wide distribution of the monitor for use on many computing systems. Due to limitations in time, manpower, and money, reprogramming on a system-by-system basis was rejected as a viable method of implementing CADA. Since no entirely transportable language

for all interactive systems existed, it was decided to pursue a strategy which would permit interdialect translation rather than actual reprogramming. Examination of available hardware and software pointed toward the BASIC programming language as the only possibility for translatability across several manufacturers. A study was then made by Isaacs (1972) which showed that programs written in one dialect of BASIC could <u>easily</u> be translated into that of many other manufacturers' dialects provided certain specified constraints on the initial programs were observed. The first BASIC version of CADA was then written by Isaacs and Christ in the BASICX dialect for the CDC 3600 at The University of Massachusetts. This was then easily and quickly translated into versions for the Hewlett-Packard 2000C and the Digital Equipment Corporation PDP-11, thus validating the assertions made by Isaacs.

The detailed outline of the current monitor was developed based on considerations falling in three basic areas--user interaction, systems constraints, and programming considerations. The user interaction is by far the most important consideration. Although the user may be highly skilled in his own subject area, he may be quite unsophisticated in terms of computer skills. The first design rule was then that the user be required to have no programming skills. He need know only three system-related commands: (1) how to sign on the system; (2) how to start the monitor running; and (3) how to sign off the system.

The second design rule was that the monitor be self-documenting in terms of options available. The monitor should be modifiable to include new models, new techniques, and improvements to current programs without the user having to wonder whether he has the latest "newsletter" or update sheet.

The third design rule was that the user should not be left "hanging". If a numerical integration fails to converge, an error message followed by the stopping of the program is not enough. Control must branch to a point where the unsophisticated user can proceed on the information available to him. Furthermore, whenever possible, input from the user must be checked for validity to avoid system errors such as division by zero, taking the root of a negative number, etc.

The constraints of any language implementation limit what can be programmed in that language. When programming for translatability across several systems, the constraints become somewhat more demanding and at times preclude the use of features that may be present on one system only, or that differ radically from one system to the next. This, with the three design rules mentioned above, has governed most of the design of the monitor and the programs.

While the monitor is currently available for operation on only three systems, an attempt has been made to minimize the dependence on features not available in BASIC dialects for other computers. The two features used which might be the most limiting are chaining and formatted print statements. However, the systems in which we are <u>most</u> interested have these features available. The formatted print statements were used to present the output and textual material in a visually pleasing way. This is not necessary, per se, but is desirable to facilitate the man-machine interaction since the intended user is not presumed to be a computer expert. The formatted print statements do have analogs in the other dialects we propose to use; however, they will be the ones needing the most change from machine to machine.

Chaining, which is necessary in some larger machines and most smaller machines, is much more central to the logical design of the system. The first consideration was that the user need only know how to sign on the system and would not need to know the names of the individual routines. This implies either a main routine-subroutine system or a monitor program which causes the loading of the proper program. The latter is the system used by us, dictated by the design of most BASIC systems. The main routine-subroutine system has the advantage of ease of parameter passing. However, the number of parameters to be passed in our system is few and the values are values known to the user, usually understood by him, and normally recorded, to be used in any published record of the analysis; thus, it is reasonable to ask the user to reenter the parameters when necessary. This also allows the user to easily do an analysis in steps at different times. The chaining as used here has the advantage of having in core only the program in use and thus reducing system overhead. A second consideration for the system is that it should be expandable with little effort on the part of the programmer and with no operational change visible to the user. The monitor system used here permits this. The only change seen by the user is that he is given the choice of choosing among a larger set of routines and techniques. The programmer need add only about three lines of coding to the monitor to make a new routine available to the user. A third consideration is that the user should never be left dangling after he makes an error. In the CADA monitor, when a program fails, the system chains to a routine in which the user ia told to save the output for use by the person maintaining the system and is then returned to the monitor to continue the session if he so wishes. All user input is screened for validity. Since string

handling capability is not highly developed in all BASIC dialects and handling a finite set of responses can be done by much simpler coding, user responses to questions within the program segments have been forced to numeric form.

Programming ease was also considered. A modular method was used in building the routines themselves. Many routines were common across programs (e.g., integrating a beta distribution, calculating an inverse chi highest density region) and were assigned specific line numbers above 5000. These routines were coded only once and after being debugged were usable without further effort on the part of the programmer. The programmer then referenced these routines by GOSUB statements to predetermined line numbers with no need to worry about where to put them. Unique portions of programs were then programmed with line numbers below 2000. As noted above, the monitor system used enables new programs to be added with little programming effort.

The accompanying appendices show a sample of the monitor output, give a listing of the current package contents, and outline the chaining sequence.

APPENDIX I

Monitor Output

```
RUN CBCADA

              COMPUTER ASSISTED DATA ANALYSIS

IF YOU WISH AN EXPLANATION TYPE 1, ELSE TYPE 0
? 1


    THIS PACKET OF PROGRAMS PROVIDES A GROUNDING IN THE
FUNDAMENTALS OF BAYESIAN METHODS OF STATISTICAL INFERENCE.
THESE ROUTINES ARE DESIGNED TO GUIDE THE RESEARCHER WHO HAS
ONLY A MINIMAL ACQUAINTENCE WITH BAYESIAN METHODS, STEP-BY-
STEP THROUGH A COMPLETE BAYESIAN ANALYSIS.  A LIST OF THE
ROUTINES FOLLOWS:

1.  PRIOR BETA-BINOMIAL MODEL
2.  POSTERIOR BETA-BINOMIAL MODEL
3.  PRIOR TWO PARAMETER NORMAL--MARGINAL DIST FOR STANDARD DEV
4.  PRIOR TWO PARAMETER NORMAL--CONDITIONAL DIST FOR MEAN
5.  POSTERIOR TWO PARAMETER NORMAL
6.  PRIOR M-GROUP PROPORTIONS
7.  POSTERIOR M-GROUP PROPORTIONS
8.  EVALUATE STUDENT-DISTRIBUTION
9.  EVALUATE BETA-DISTRIBUTION
10. EVALUATE INVERSE CHI-DISTRIBUTION
11. EVALUATE NORMAL DISTRIBUTION
14. CALCULATE MEANS,STANDARD DEV., SUMS OF SQUARES

IF YOU WANT TO RUN ONE OF THE ABOVE ROUTINES,TYPE ITS NUMBER
OTHERWISE TYPE A ZERO.
? 1
```

APPENDIX II

Package Contents

I. Supervisory Routines

  A. CADA - Monitor

  B. ERROR - Gives instructions when a program fails

II. BETA - Binomial Model Routines

  A. PRIORB - Assists in fitting prior knowledge to the beta class

  B. POSTB - Combines a beta class prior with binary data to give a beta posterior

III. Two Parameter Normal Model

  A. PRIORS - Fits prior knowledge (marginal) on the standard deviation to an inverse chi distribution

  B. PRIORM - Fits prior knowledge (conditional) on the mean to a normal distribution

  C. POSTN - Combines the inverse chi and normal priors with normal data to give posterior distribution

IV. m-Group Proportions

  A. PRIORP - Evaluates exchangeable prior information on any of a set of proportions for use in an m-group proportion routine

  B. PROPOR - Solves the Lindley equations for a set of binary data

V. Evaluation Routines

  A. TDIST - Evaluates the probability integral of a nonstandard student t-distribution

  B. BDIST - Evaluates the probability integral of a beta distribution

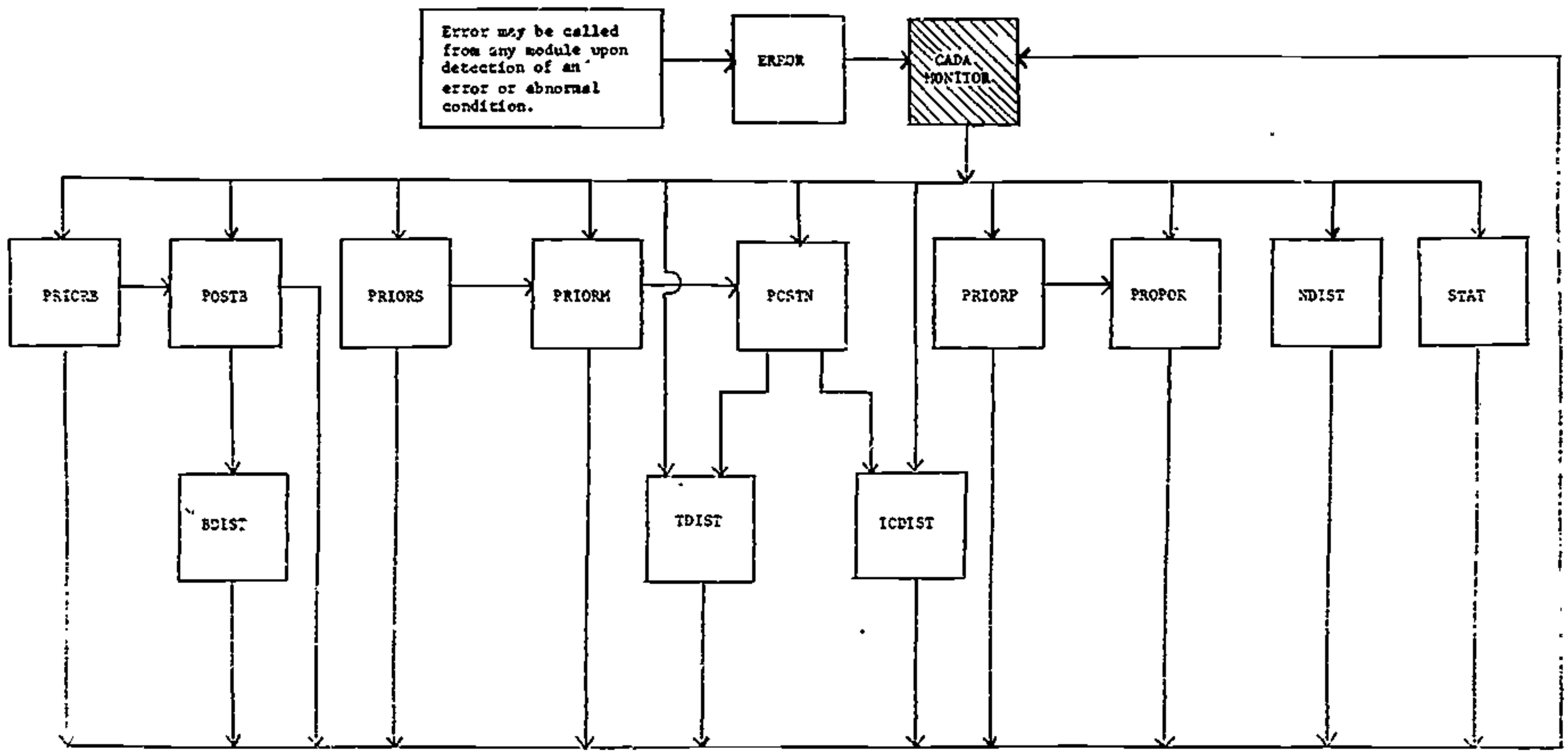
C. ICDIST - Evaluates the probability integral of a nonstandard inverse chi distribution

D. NDIST - Evaluates the probability integral of a nonstandard normal distribution

VI. Service routine STAT calculates the mean, standard deviation, and sum of squared deviations from the mean for a set of data

APPENDIX III

Chaining Sequence



Note: Any program can chain to error upon detection of an abnormal condition.

Bibliography

Isaacs, G. L. Interdialect translatability of the BASIC programming language. ACT Technical Bulletin No. 11. Iowa City, Iowa: The American College Testing Program, 1972.

Novick, M. R. Computer-assisted data analysis. ACT Technical Bulletin No. 3. Iowa City, Iowa: The American College Testing Program, 1971.